量子生物学及其应用

刘次全 主编
江寿平 审校

高等教育出版社

58.1713
191

量子生物学及其应用

刘次全　主编
江寿平　审校

中国科学院植物研究所图书馆藏

高等教育出版社

25650
中科院植物所图书馆
S0014731

内 容 提 要

本书内容丰富、广泛，反映了近代量子生物学的主要方面及最新研究成果和发展趋势。全书共12章，分别介绍了量子生物学的基本原理和方法，以及应用量子生物学方法对核酸结构与功能、蛋白质电子结构、酶作用机理、生物分子空间构象、金属离子生物学功能、生物系统中能量和电子转移、化学致癌作用等的研究，并介绍了量子遗传学及量子药理学。

本书适合生物学、医学、农学及物理、化学等专业的研究生和大学高年级学生使用，对教师和有关科研工作者也是一本有价值的参考书和工具书。

编著者（以姓氏笔划为序）

王　莹　（中国科学院昆明动物研究所）
王身立　（湖南师范大学）
白逦彬　（中国科学院生态研究中心）
陈　洪　（北京轻工业学院）
吴吉安　（中国科学院上海药物研究所）
刘次全　（中国科学院昆明动物研究所）
刘征先　（中国科学院上海冶金研究所）
沈仲钧　（华东师范大学）
邵　俊　（上海科学技术大学）
曹　槐　（云南大学）
谭载友　（广东医药学院）
中国科学院上海生物化学研究所生物大分子构象组

责任编辑　田　年

量子生物学及其应用
刘次全　主编
江寿平　审校
*
高等教育出版社出版
新华书店总店北京科技发行所发行
人民教育出版社印刷厂印装
*
开本 787×1092 1/16　印张 36.75　字数 860 000
1990年12月第1版　1990年12月第1次印刷
印数0001— 1 220
ISBN 7-04-003216-3/Q·172
定价 7.20 元

目　录

前　言

第1章　绪　论

1.1　一门新兴的边缘学科——量子生物学 ……（1）
1.2　量子生物学的研究内容 ……（2）
1.3　量子生物学的任务 ……（2）
1　量子生物学与其它学科的联系 ……（3）
1.5　量子生物学展望 ……（5）

第2章　量子生物学的理论方法和理论

2.1　量子化学基础知识 ……（7）
2.1.1　波函数与轨道 ……（7）
2.1.1.1　波函数 ……（7）
2.1.1.2　轨道 ……（8）
2.1.2　量子化学中的基本近似 ……（10）
2.1.2.1　非相对论近似 ……（10）
2.1.2.2　Born-Oppenheimer 近似 ……（10）
2.1.2.3　轨道近似 ……（11）
2.2　量子生物学方法和指数 ……（12）
2.2.1　简单分子轨道法 ……（12）
2.2.1.1　Hückel 分子轨道法（HMO 法） ……（13）
2.2.1.2　微扰分子轨道法（PMO 法） ……（20）
2.2.1.3　扩展的 Hückel 分子轨道法（EHMO 法或 EHT 法） ……（23）
2.2.2　半经验自洽场分子轨道法 ……（26）
2.2.2.1　全略微分重叠法（CNDO 法） ……（27）
2.2.2.2　修改的忽略双原子微分重叠法（MNDO 法） ……（31）
2.2.3　从头计算法（ab initio 法） ……（33）
2.2.4　X_α 方法 ……（38）
2.2.5　分子静电势法 ……（40）
2.2.5.1　从头计算法的分子静电势计算 ……（41）
2.2.5.2　CNDO 法分子静电势计算 ……（44）
2.2.5.3　生物分子的分子静电势计算 ……（46）
2.2.6　PCILO 计算法 ……（46）
2.2.6.1　PCILO 法基本原理 ……（47）
2.2.6.2　PCILO 法的一般概要 ……（49）
2.2.6.3　PCILO 法的适用范围以及应用实例 ……（50）
2.2.7　分子力学计算法 ……（52）
2.2.7.1　分子力学原理 ……（52）
2.2.7.2　分子力学计算流程 ……（56）
2.2.7.3　分子力学在量子生物学中的应用 ……（56）

第3章　原子间相互作用和分子间相互作用

3.1　生物分子中原子间的相互作用 ……（61）
3.1.1　相互作用力和能量 ……（61）
3.1.2　原子间相互作用的分类 ……（63）
3.1.3　定域键和离域键 ……（63）
3.1.3.1　定域键 ……（63）
3.1.3.2　离域键 ……（67）
3.2　分子间的相互作用 ……（69）
3.2.1　Van der Waals-London 相互作用 ……（70）
3.2.2　分子间的氢键相互作用 ……（75）
3.2.3　电荷迁移相互作用 ……（79）
3.3　生物分子的特异作用和识别 ……（84）
3.3.1　特异作用和生物分子识别的物理、化学基础 ……（84）
3.3.2　酶与底物的专一作用 ……（87）
3.3.2.1　活性中心与诱导-契合学说 ……（87）
3.3.2.2　酶专一性本质的进一步探讨 ……（88）
3.3.2.3　从轨道对称性看酶的专一性 ……（89）
3.3.2.4　酶与辅酶间的相互作用 ……（89）
3.3.3　密码子与反密码子的识别和

相互作用 ………………………………（92）
3.3.4 抗原-抗体反应的专一性 …………（95）

第4章 核酸的结构与功能

4.1 核酸的化学组成和分子结构 ……………………………（99）
4.2 核酸及其组分的电子结构………（104）
4.2.1 嘌呤和嘧啶碱基的电子结构 ……（105）
4.2.1.1 电荷分布和偶极矩 ……………（105）
4.2.1.2 碱性 …………………………（109）
4.2.1.3 分子轨道能、电离势和电子亲合势 ……………………………（110）
4.2.1.4 电荷密度、自由价、键级和超离域度 ……………………………（113）
4.2.1.5 碱基激发态的电子结构 ………（115）
4.2.2 碱基的互变异构体 ……………（118）
4.2.3 碱基的分子静电势 ……………（121）
4.2.4 核酸碱基的配对 ………………（121）
4.2.5 多聚核苷酸 ……………………（126）
4.2.5.1 核苷 …………………………（126）
4.2.5.2 核苷酸 ………………………（126）
4.2.5.3 多聚核苷酸 …………………（131）
4.3 DNA 双螺旋构象的稳定性……（134）
4.4 核酸的辐射效应和光生物学问题 ……………………………（136）
4.5 核酸的能带结构 ………………（139）
4.6 核酸和蛋白质的相互作用………（142）
4.6.1 碱基对与丝氨酸的专一识别 ……（142）
4.6.2 B-DNA与多肽的相互作用………（143）
4.6.3 遗传密码子与氨基酸的相互作用…………………………………（145）
4.7 核酸和水的相互作用 …………（146）
4.8 核酸的分子静电势研究 ………（149）
4.8.1 核酸基本组成单元的分子静电势…………………………………（149）
4.8.2 较高单元的分子静电势 ………（150）
4.8.3 一段完整的 B-DNA 双螺旋的分子静电势 ……………………（151）
4.9 核酸的功能 ……………………（153）
4.9.1 作为遗传信息载体的 DNA ……（153）
4.9.2 核酸中电子激发能的转移方式 …………………………………（156）
4.9.3 DNA的复制……………………（158）
4.9.4 DNA的甲基化作用与基因功能…（160）

第5章 蛋白质的电子结构

5.1 氨基酸的电子结构和氨基酸溶液结构的模拟…………………（166）
5.2 短杆菌肽A离子选择性的蒙特卡洛研究 ……………………（183）
5.3 肽链构型的能量分析 …………（191）
5.3.1 肽结构的描述 …………………（191）
5.3.2 规则多肽链分子构型的能量分析……………………………………（193）
5.3.3 聚 L-丙氨酸的 ab initio 计算 ……………………………………（199）
5.4 蛋白质分子的量子化学研究 ……………………………………（203）
5.4.1 胰岛素的量子化学研究 ………（204）
5.4.1.1 胰岛素分子的疏水内核和分子的几何特征 …………………（204）
5.4.1.2 胰岛素分子静电势的计算……（208）
5.4.1.3 胰岛素分子相互作用能的计算……（213）
5.4.1.4 胰岛素分子主链的 CNDO/2 计算 ……………………………（219）
5.4.2 木瓜蛋白酶的量子化学研究 ……（219）
5.5 蛋白质的能带结构 ……………（231）

第6章 酶作用机理的量子生物学研究

6.1 模型和方法……………………（242）
6.1.1 酶反应模型 ……………………（243）
6.1.2 催化部位的分子轨道计算 ………（244）
6.1.3 分子力学计算 …………………（246）
6.1.4 轨道取向、轨道对称性和稳定能 ………………………………（246）
6.1.5 环境的模拟 ……………………（250）
6.2 羧肽酶A和其它金属酶的量子生物学研究 …………………（253）

第7章 量子遗传学

7.1 量子遗传学的背景………………（263）

7.1.1 从经典遗传学到量子遗传学 …… (263)
7.1.2 分子遗传学概要 …… (267)
7.1.2.1 基因及其复制 …… (267)
7.1.2.2 基因及其表达 …… (268)
7.1.2.3 基因表达的调控 …… (271)
7.2 原子和电子水平的信息流 …… (273)
7.2.1 基因复制和突变的微观机制 …… (273)
7.2.1.1 遗传密码的基本结构 …… (273)
7.2.1.2 DNAY型复制的复制平面模型 …… (275)
7.2.1.3 遗传密码的质子-电子对模型 …… (277)
7.2.1.4 一级微扰模型 …… (284)
7.2.2 基因表达及其控制的数理模型 …… (290)
7.2.2.1 DNA 转录的物理模型 …… (290)
7.2.2.2 密码子 …… (291)
7.2.2.3 转译模型 …… (301)
7.3 量子遗传学现状 …… (304)

第8章 量子药理学

8.1 引言 …… (313)
8.2 药物分子构象 …… (315)
8.2.1 构象对药物分子药效的重要性 …… (315)
8.2.2 分子轨道法在研究药物分子构象中的应用 …… (318)
8.2.3 去甲肾上腺素的构象研究 …… (323)
8.2.4 多巴胺构象的研究 …… (325)
8.2.5 抗生素类抗肿瘤药物的量子药理学研究 …… (327)
8.2.6 作用于中枢神经系统药物药效构象的研究 …… (332)
8.3 药物的受体 …… (335)
8.3.1 受体学说在药物研究中的重要性 …… (335)
8.3.2 乙酰胆碱的量子药理学研究 …… (340)
8.3.3 5-羟色胺受体 …… (342)
8.3.4 酶抑制剂 …… (346)
8.3.5 受体亚型 …… (350)
8.4 药物分子的静电势和电荷密度 …… (354)
8.4.1 量子药理学中静电势和电荷密度的计算 …… (354)
8.4.2 组胺及其衍生物的电荷密度和分子静电势 …… (356)
8.4.3 β受体药物的分子静电势研究 …… (359)
8.4.4 增敏剂的量子药理学研究 …… (362)
8.4.5 苯甲酰胺型 D_2 受体拮抗剂的分子静电势计算 …… (366)
8.5 电荷转移 …… (368)
8.5.1 电荷转移中的量子药理学研究 …… (368)
8.5.2 致幻剂的电荷转移研究 …… (372)
8.5.3 抗心律失常新药常咯啉及其类似物的量子药理学研究 …… (374)
8.5.4 镇痛药中的电荷转移研究 …… (376)
8.5.5 吗啡类药物的电荷转移复合研究 …… (377)
8.6 药物分子设计与 QSAR …… (380)
8.6.1 引言 …… (380)
8.6.2 QSAR 中 Hansch 途径的基础 …… (382)
8.6.3 QSAR 研究中的参数 …… (384)
8.6.4 QSAR研究的应用 …… (386)
8.6.5 其他药物分子设计方法 …… (388)

第9章 生物分子的空间构象

9.1 蛋白质和核酸分子二级结构的预测 …… (392)
9.1.1 蛋白质分子二级结构的预测 …… (392)
9.1.1.1 Chou和Fasman方法 …… (393)
9.1.1.2 各种预测方法结果的比较 …… (396)
9.1.1.3 蛋白质二级结构预测的计算机程序 …… (397)
9.1.2 核酸分子二级结构的预测 …… (404)
9.1.2.1 RNA 大分子的二级结构预测方法 …… (405)
9.1.2.2 5.8SrRNA的二级结构 …… (409)
9.1.2.3 大肠杆菌RNA聚合酶识别的启动子的结构特征 …… (413)
9.2 蛋白质和核酸的构象和相互作用 …… (417)

9.2.1 势能函数 …… (417)
9.2.2 分子间相互作用 …… (419)
9.2.2.1 甲酰氨-水的相互作用 …… (420)
9.2.2.2 核酸碱基分子的水合作用 …… (421)
9.2.2.3 阿霉素与核酸的相互作用 …… (422)
9.2.3 核苷酸和蛋白质空间构象和电子结构的计算 …… (423)
9.2.3.1 经验势能法 …… (423)
9.2.3.2 分子轨道方法 …… (430)
9.3 分子动态学 …… (436)
9.3.1 蛋白质动态学 …… (436)
9.3.2 酶-底物反应动力学 …… (439)

第10章 金属离子生物学功能的量子生物学研究

10.1 概述 …… (443)
10.2 金属离子生物学功能的配位场理论 …… (443)
10.2.1 与金属离子相互作用的生物配位体 …… (444)
10.2.2 配合物的晶体场理论 …… (444)
10.2.3 配合物的分子轨道理论 …… (447)
10.2.4 配位场理论的应用 …… (448)
10.3 金属离子与核酸间的相互作用 …… (449)
10.3.1 金属离子与核酸碱基相互作用的最优配位模式 …… (449)
10.3.2 金属离子对核酸构象稳定性的影响 …… (455)
10.4 金属离子与蛋白质的相互作用 …… (460)
10.4.1 金属离子对蛋白质能带结构的影响 …… (461)
10.4.2 金属的螯合性质对酶活性的影响 …… (464)
10.5 金属-卟啉配合物的分子轨道研究 …… (468)
10.5.1 卟啉配合物的分子轨道计算(一) …… (469)
10.5.1.1 铁卟吩分子轨道的构造 …… (470)
10.5.1.2 能级的排布 …… (470)
10.5.1.3 电荷分布 …… (472)
10.5.2 卟啉配合物的分子轨道计算(二) …… (473)
10.6 金属离子生物学功能的一些量子生物学观点 …… (477)
10.6.1 金属离子与生物大分子相互作用的若干问题 …… (477)
10.6.1.1 半导体掺杂观点 …… (477)
10.6.1.2 电子-构象相互作用观点 …… (478)
10.6.2 金属离子与氧的积累、贮存和输运 …… (482)

第11章 生物系统中能量和电子的转移

11.1 生物大分子的能量转移 …… (489)
11.1.1 能量的共振转移 …… (490)
11.1.2 激子转移 …… (491)
11.1.3 孤子转移 …… (495)
11.1.3.1 一维弹性分子结构中的孤子 …… (496)
11.1.3.2 激子与孤子性质的比较 …… (499)
11.1.3.3 α-螺旋蛋白质中分子结合能沿一维分子链传输的孤子运动 …… (499)
11.2 生物系统中的电子转移 …… (502)
11.2.1 络合物电荷转移 …… (503)
11.2.2 电子的隧道转移 …… (505)
11.2.3 蛋白质在电子远距离转移过程中的作用 …… (506)
11.2.4 质子转移与氢键 …… (512)

第12章 化学致癌作用的量子生物学研究

12.1 概述 …… (515)
12.1.1 历史 …… (515)
12.1.2 化学致癌物的发现 …… (515)
12.1.3 体内、体外代谢实验与最终致癌物 …… (516)
12.1.3.1 代谢途径 …… (516)
12.1.3.2 最终致癌物 …… (517)
12.2 结构与活性关系研究 …… (518)
12.2.1 多环芳烃(PAH) …… (518)
12.2.1.1 L-区理论 …… (518)

12.2.1.2 K-区理论 …………………………… (522)
12.2.1.3 湾区理论 ………………………………… (522)
12.2.1.4 双区理论 ………………………………… (523)
12.2.1.5 多环芳烃的构型效应 ……………… (526)
12.2.1.6 多环芳烃致癌能力的参数图 …… (535)
12.2.1.7 湾区 π 键级、分子大小与致癌能力关系 ………………………… (537)
12.2.1.8 多环芳烃的定量结构致癌活性关系 ……………………………………… (538)
12.2.2 甲基多环芳烃 ……………………… (541)
12.2.2.1 甲基的空间效应和电子效应 … (541)
12.2.2.2 甲基取代多环芳烃的致癌活性 ………………………………… (543)
12.2.3 亚硝胺类……………………………… (550)
12.2.4 微生物毒素………………………… (551)
12.2.5 多环芳烃“双区”理论的发展…… (553)
12.2.5.1 多环芳烃“双区”理论的定量研究 ……………………………………… (553)
12.2.5.2 甲基、烷基代多环芳烃等的定量结构致癌活性关系的研究 …… (553)
12.2.5.3 多环芳烃的致癌性与其分配系数之间的定量关系 ……………… (556)
12.3 化学致癌物同生物大分子的相互作用 ……………………………… (556)
12.3.1 同核酸的相互作用 ………………… (557)
12.3.2 同蛋白质的相互作用 …………… (561)
12.3.3 致癌物作用后大分子的改变…… (562)
12.4 化学致癌作用的量子生物学研究方法 …………………………… (563)
12.4.1 量子化学方法 …………………… (563)
12.4.1.1 孤立分子近似 ……………………… (564)
12.4.1.2 位能面 ……………………………… (567)
12.4.1.3 分子静电势 ………………………… (567)
12.4.2 多元回归分析 …………………… (570)
12.4.3 模式识别…………………………… (571)
12.4.3.1 模式识别法简介 ………………… (571)
12.4.3.2 模式识别在化学致癌作用研究中的应用实例 …………………… (574)

前　　言

生命科学在近三十年来的飞跃进展令人振奋。不单是生物学家，还有数学家、物理学家、化学家、工程师等都以其本门学科的专业知识和特有技术，参与生物学的研究，从各个角度揭露生命科学的奥秘。其中参与研究的量子生物学是一支生力军，这门学科的特点是处在生物学、物理学、化学、数学等几门不同学科的接触与交汇点。

量子生物学运用量子力学原理，通过数学运算来研究生物分子内的电子运动。生物科学可以通过不同层次同时进行研究，例如细胞生物学从细胞层次进行研究，分子生物学从分子层次进行研究，而量子生物学则从原子(电子)的层次进行研究，从电子水平来理解生命现象。参与生命过程的分子或原子间作用，其中包括化学反应以及原子间作用的力(例如氢键作用等)都离不开电子运动。从这意义说来，研究生命过程的电子运动正是从本质上探索生命现象。

量子生物学目前还只是一门新兴学科，仅在近几十年来量子力学不断发展后，才逐渐和分子生物学结合而形成量子生物学。可是，从30年代Pullman所著量子生物学一书问世以来，量子生物学已经有了很大进展，运算的方法也已大为改观。当前应用量子生物学以探索致癌机理、生物固氮、底物与酶以及神经递质、激素或药物与受体之间的识别等，已能解释若干现象的机理，并进而试行预测一些变化，说明它是一门大有前景的学科。在量子生物学的若干方面，我国科学家也已经或正在作出不少贡献。

量子生物学既是新兴学科，对许多科研工作者说来还较陌生，有加以介绍与概述的必要。刘次全等几位量子生物学专家，在教学上已积累多年经验，在科研上又卓有成果，驾轻就熟撰写《量子生物学》，叙述学科的基本内容，并且介绍国际上在该领域研究的最新进展与发展趋势，也回顾国内的研究成果，不仅可应教学之需而且可供同行参考，还可帮助其他邻近学科的从业人员了解该领域的概况。

当然，生命科学的探索，必须分别从不同方面、不同角度、不同方法进行研究，其中量子生物学仅是从理论上进行数学运算，只是方法之一。本书既然专门介绍量子生物学，不免加以强调，但这绝不意味着排斥其他层次、其他角度、其他方法的研究。相反，量子生物学既是新兴学科，并非完满无缺，更必须与其他学科，尤其实验性学科配合，才能向理解生命的深处迈进！

嵇汝运

1987年10月

于中国科学院上海药物研究所

第1章 绪 论

1.1 一门新兴的边缘学科——量子生物学

一切自然科学的发展几乎都要经历观察、实验、描述和理论各阶段。作为研究生命现象的生物学,也必然要经历科学发展的共同道路。

纵观生物学的发展历史可以看出,近代生物学已经从生物圈、生态、群落、种群、个体、系统、器官、组织和细胞的研究层次,深入到了分子和电子的层次。这既是生物学发展的必然趋势,也是其它自然科学和技术科学向生物学渗透的结果。今天的生物学已经发展成为多研究层次的、具有"横跨性"的综合学科,它正处于并将继续处于无比活跃和引人注目的时期。

须要着重强调的是,在物理学、化学、数学、计算机科学和实验技术飞速进步的形势下,自50年代以来,出现了分子生物学急剧发展并取得重大成果的时期,从而推动了生命科学由宏观到微观,由观察、实验到定量化、理论化的进程。

在分子生物学发展的同时,有人主张从更微观的结构,即从电子一级水平来解释生命现象和研究生命过程,这被称之为量子生物学。事实上,从50年代起,运用电子理论来解释和阐明生物分子微观结构和运动规律,以及某些生命过程的研究工作就已经逐步开展起来了。然而,量子生物学真正取得国际科学界的重视还只是十几年前的事情。在生物学家的呼吁和量子物理、量子化学家们的共同努力和积极参与下,作为一门学科领域形成标志的学科组织也相继诞生了。1970年,由量子物理学家 Lõwdin 和分子药理学家 Purcell 等倡议成立了国际量子生物学会(简称 ISQB) 。

量子生物学是一门研究生命物质微观结构的理论性学科,它是量子力学与分子生物学相结合的产物。它运用量子力学原理,通过数学运算来研究生物分子的电子结构、电磁性质、能量转移和化学反应等问题。

量子生物学的发展历史不长,尚处于幼年时期。然而,它已经在化学致癌机理、生物固氮、病理机制的分子电子基础、抗原-抗体作用的分子识别(以及酶与底物、药物与受体和碱基配对等的分子识别)、金属离子在生命活动中的作用、药物分子作用的电子机理、生物膜的研究以致可以逐步实现"药物分子设计"等等方面都已经取得或正在取得一系列进展,因而在客观上,引起了人们的日益注视。但是,与此同时,我们必须看到,量子生物学还有许多不完备的地方,因而错误也在所难免。更何况量子力学本身还存在着困难和不足呢!在这点上,Szent-Györgyi 的下列言论给了我们有益的启示:"只有一条道路能够避免错误,那就是什么事情也不做,或者避开一切新事物。但是,它却是在所有这些错误中最大的错误"[1]。为了减少和克服量子生物学研究中迄今还存在的某些不足,使我们的理论计算能够直接与生物学中的经验事实相联系,强调生物机体是一个高度有序并与环境保持协调的复杂体系;强调在生物体中分子与分子之间存在着复杂的相

互作用；强调生物分子的多维结构和动态特征；强调分析与综合的研究思想和方法；强调计算方法和计算结果的局限性是十分必要的。在这样的前提下，理论计算才可能是有价值的。迄今为止，我们已经可以举出若干例子来说明理论计算有时甚至可以比实验得出更为精确细致的分子性质(这将会在以后的有关章节中讲到)，尤其是当实验不能进行时，在这样的情况下，计算就更为有用。

由于量子生物学是一门边缘性的新兴学科，因而需要综合有关学科的知识、技术和成就，以及实验和理论研究工作者的密切配合，才可望取得更多的实质性的进展。

以上我们就量子生物学的意义和发展历史作了简略的介绍。那末，到底什么是量子生物学呢？(这也是一些生物学家经常提到的问题)不同的人给予了不同的定义。为统一定义，国际量子生物学会曾经讨论过这个问题并作出了如下结论："量子生物学即是量子力学作为工具在生物学问题上的应用"。由于量子力学是处理基本粒子，特别是电子行为的学问，因此，换句话说量子生物学就是从电子水平来理解生命现象的学问。

1.2 量子生物学的研究内容[2]

从量子生物学当前的研究内容看，它将过去被称为"亚分子生物学*"、"电子生物学"、"电子生物化学"、"量子生物物理学"、"量子生物化学"、"量子药理学"以及"生物能力学"等的部分内容，都包括在其研究范围之内。就发展趋势看，作者认为，量子生物学可望向着实验和应用方向发展，向着与分子生物学，乃至细胞生物学紧密联系的方向发展，因而其研究内容也将会不断更新。目前和今后的一段时间内，量子生物学的研究内容大致有以下几个方面：

生物分子的电子结构与反应活性的关系；

生物分子间的相互作用；

生物大分子的构象及溶剂对构象的影响；

特异相互作用与识别；

致癌物质的结构与致癌活性；

生物体中的电子转移、质子转移、能量转移；

无机离子在生物过程中的作用；

遗传、突变的量子理论；

量子药理学等。

1.3 量子生物学的任务

在当今生物学中，分子水平的研究已经取得并将继续取得重大成果。然而，分子生物学发展的同时又告诉我们，生命过程在电子层次的研究也是不可缺少的。

---

* M. B. Вопькенштейн 认为"亚分子生物学"的说法是毫无意义的[3]。

我们知道，物质的宏观性质是多种多样的，而决定这些性质的原因也是多种多样的，不可一概而论。可是，通过科学实验，人们已经知道有一大批性质，包括几乎全部化学性质、半导性、电磁性质和许多光学性质，以及与生命的基本过程有关的性质或生物功能，都跟物质的分子结构，尤其是组成分子的原子外层电子的运动有关。事实说明，诸如遗传、变异、免疫、细胞功能、生长、发育和繁殖等等生命现象，都与生命的物质基础——蛋白质和核酸等生物分子的特定结构(包括电子结构)和运动规律有着密切的关系。人们通过不断地研究自然界的生物机体后认识到，生命的奥秘部分地存在于微观世界之中。例如某些生命缺陷(如致癌、突变)只是从群体、个体、系统、器官、组织、细胞乃至分子水平上加以认识，往往是很不够的，而必须配合原子或电子水平的研究，才能真正弄清其本质。这可以看成是对生物大分子进行量子力学处理和计算的必要性。

今天，分子生物学的成就已经显示，有些基本生命现象需要在分子水平上予以描述。但要回答诸如：

(1) DNA 双螺旋中，A—T、G—C 形成特异的氢键，这些氢键是由怎样的力形成的？

(2) 决定氨基酸三联体密码的因子到底是什么？为什么色氨酸只对应于 UGG 而不是其它的三联体？

(3) 双链 B-DNA 或者蛋白质的 α 螺旋为什么是右旋的，而双链 Z-DNA 又是左旋的？

(4) 生物分子 (酶与底物，抗原与抗体，药物与受体，DNA 与解链蛋白等等)是依靠什么识别的？

(5) 为什么酶蛋白和其它生物催化剂有很高的催化效率？

(6) 蛋白质和核酸是否具有半导性？

(7) ATP 为什么包含有高能键？

(8) 生物系统中能量传递和电荷传递间的关系怎样？

(9) 遗传信息流在电子水平上的传递机理，等等。

以上这些问题涉及到分子中的电子，而分子生物学是无法描述生物分子中运动着的电子的。可是，量子生物学的理论计算却能给出有关电子分布的更多的细节。譬如了解得最清楚的 B-DNA 的立体结构已经给出了分子中原子核的空间位置，在 Born-Oppenheimer 近似下，原子核和电子可以处理成各自分离的运动。这就提供了进行生物大分子电子结构的量子力学计算的可能性。迄今为止，量子生物学在一些问题上所取得的成就，正好说明生命现象在电子水平上描述的必要性和可能性。

综上所述可见，量子生物学的任务即是在电子水平上描述和探索生命现象。

1.4 量子生物学与其它学科的联系[4]

既然量子生物学是一门新兴的边缘科学，那么它与其它一些学科(如物理学、化学和数学等)的关系无疑是十分密切的。量子生物学作为近代生物学的一个新的分支，又有它固有的特征。

在自然科学发展的早期，就已经显示出生物学与物理学的联系以及由于这样的联系而获得

的一系列成果。从现代观点看，生命是复杂分子系统通过跟其它系统交换能量、物质和信息而发生的特殊物理和化学过程。

尽管生物体是形形色色的，但构成他们的化学单位并不特殊。譬如所有的蛋白质都是由大约 20 种氨基酸按不同的数量和排列顺序所组成的多聚物，而 DNA 则是由四种脱氧核糖核苷酸组成的多聚物。至于原子和组成原子的基本粒子乃是物质的基本要素。为此，我们可以认为，处理非生命物质的原子、分子、分子间的相互作用、能量、电子和质子及其转移等的量子力学、统计物理学和热力学方法(特别是非平衡态热力学)，一般地讲，同样可以应用于生物分子。

在量子理论中一次计算波动方程的解就能得出有关原子、分子和某些更大体系的较多的结构性质和与之有关的信息，因而有可能对这些系统的行为提供较为全面的了解。将量子力学应用到生物学问题上还有许多优点，那是因为量子力学能以相当直接的方式告诉我们生物分子内的电荷分布。当我们欲寻求一种活性分子在它起主要作用的瞬间，所呈现的精确的三维电子密度究竟是怎样的？那么量子力学则是提供这一答案的最好手段。

事实说明，在生物分子功能的研究中，量子力学发挥着重要的作用。在光合作用和视觉原初过程的研究中；在遗传突变和 DNA 分子之间关系的研究中；在分析酶反应中底物的特异性和反应的特异性时，量子力学的处理方法具有特殊的意义。

作为量子生物学基础的量子物理学并不是一门停滞的学科，它也正在经历着变革与发展。诚然，量子力学本身直至今天为止，它的理论基础还有许多可争论之处，如波函数的统计解释，应用到场上产生的发散困难等等。尽管如此，由它所导出的结论却与相当量的实验事实符合得很好，这也正是它所取得的成功之处。当然，我们必须认识到，生命现象毕竟具有一系列独特的不同于非生命的特征。首先，生命运动属于更高一级运动形式。生物体是高度有序的、复杂的、多相的。同时，生物分子之间又存在相互作用。

生物学与化学的联系也是由来已久的，而量子生物学与量子化学的联系就更为紧密。事实上，就方法学而论，量子生物学的理论计算方法大多是沿袭量子化学的方法(这里所指的是理论计算方法)。但是，不能因此认为，量子化学可以代替量子生物学(作者不同意将量子生物学看作是化学学科的一个分支的观点，其理由已在上面的讨论中强调过了)。量子生物学研究的对象是远比量子化学研究的分子大得多的、且更复杂的分子。何况对前者来说，重要的是要在静态结构基础上研究生物分子的动态过程以及生物分子的结构与功能受环境影响的程度，因而困难更大，要求在方法学上作进一步的改进和发展。

在生物学向定量化、理论化发展的进程中，一方面从分析的角度发展了像分子生物学、量子生物学这样的学科，从基本的和电子的水平上去研究生命现象中的基本物理、化学过程；另一方面又从综合的、系统的角度发展了生物控制论、生物信息论和生物系统论这样一些学科，从整体的观念出发去研究复杂的生物系统。

今天，大概谁也不会否认物理和化学定律在研究生命现象中的某些适用性。大多数科学家认为，制约非生命系统的物理和化学规律有可能在最终说明生命的极其多种多样的表现。此外，也有些见解认为，需要发展新的，迄今还没有建立的物理学原理才能最终说明生命的多种表现。总之，无论这些观点的正确性如何，在实际应用物理、化学等的理论和方法探讨生物分子的运动

规律时，我们必须认识到这些理论和方法的局限性。但是，另一方面，如果能够将这些理论和方法有效地与生物学问题结合起来，充分发挥学科间相互渗透的威力，则不仅能为寻找新的生物学规律提供许多有益的启示，同时还将大大丰富和发展现有的理论和方法。即使在应用于生物学研究的过程中，发生了现有理论与生物学结果相矛盾的现象，这也是很有意义的。因为科学发展的重大突破口往往就是在矛盾的最尖锐处。

1.5 量子生物学展望

在今天，我们已经看到了量子生物学在一些领域取得了初步的，有些甚至可以认为是重大的成就。在推进量子生物学发展的进程中，我国的科学工作者做出了积极的、有相当影响的贡献。如自洽场叠代中收敛问题的研究，ATP 等较大生物分子的从头计算[6]，量子化学的计算方法和计算程序[7]，PAH 致癌的双区理论[8]，生物分子的溶液构象研究[9]等，以及本书部分作者的研究工作在国内外都占有一定的地位。

总之，量子生物学乃是一个方兴未艾、不断创新的领域，以致国内外有不少的物理学家、化学家和数学家、药理学家涉足到生物学的这一领域中来。有个哲学家曾经说过："在一门科学与另一门科学的接触点上，双方都宣称与己无关，但是恰恰就在这一点上可望取得最大的成果"。量子生物学正是处在生物学、物理学、化学和数学几个方面的共同接触点上，因此发展十分迅速，研究者非常活跃，一批基础较好并有志从事量子生物学研究的青年科学工作者正在成长。

随着量子化学计算方法的改进和发展，以及大容量高速电子计算机的问世，在量子生物学中，那些过去认为根本不可能处理的生物分子，现在有的已经实现了较为准确的量子力学计算。对于一些目前无法用从头计算方法计算，而用一般的近似方法计算也还存在着困难的生物分子，近年来，法国的 Pullman 夫妇发展了一种近似方法，即把大分子分成许多碎片，对每个碎片用从头计算方法，而碎片的连接用一般的近似计算方法。此外，还可像 Clement'e 那样计算一部分内容，其余用统计的方法，或者把大分子结构与 Thomas-Fermi 模型结合起来，计算这些生物分子(参见后续章节)。

如前所述，量子生物学是运用量子力学原理，通过数学运算来实现其研究目的的。电子计算机的出现和发展，不仅使得许多复杂的量子生物学计算得以实现，而且凭借它的记忆、判断和逻辑推理的性能，已可直接用来进行某些生物分子和药物分子的鉴别和合成路线的设计。

量子生物学研究领域十分广泛，它异军突起，表现出很强的生命力，正从一个新的角度——电子水平来研究生命过程。在分子生物学急剧发展的形势下，对生命运动在电子水平上提出的问题的阐述，将历史地落在量子生物学工作者的身上。

最后，必须指出，量子生物学还是一门比较年青的、不成熟的和发展中的学科。要使它成为研究生物分子和生命现象的有效工具，还要走一段崎岖的路程，还将会遇到种种困难和曲折。

在结束本章讨论的时候，作者以为着重强调下述观点将是有益的，即正如细胞生物学是细胞层次的生物学，分子生物学是分子层次上的生物学那样，量子生物学乃是原子(电子)层次上的生物学。在这样的意义上，量子生物学实质是在电子水平上探索生命现象的、综合性的生物学学

科。而那种将量子生物学定义为："量子生物学即是量子力学作为工具在生物学问题上的应用"，只不过是更多地从方法学角度来考虑的。

参 考 文 献

[1] 永田亲义著，陶宗晋、江寿平等译：«量子生物学入门»，上海科学技术出版社(1979)。
[2] 刘若庄、陈润生：分子科学与化学研究，(2):121(1982)。
[3] M. V. 伏尔更斯坦著，龚少明译：«现代物理学与生物学概论»，复旦大学出版社 (1985)。
[4] Давыдов, A. C.: "Биология и Квантовая Механика", (Наукова Думка, Киев), (1979).
[5] 宣建诚、江寿平：中国科学(B辑)，(7):681(1986)。
[6] 江寿平、宣建诚：中国科学(B辑)，(6):609(1987)。
[7] 吴吉安、嵇汝运：«微型计算机量子化学程序集»。
[8] 戴乾圜：中国科学，(10):964(1979)。
[9] 陈润生等：中国科学(B辑)，(1):45(1985)。

刘次全
王 莹

第2章　量子生物学的理论方法和理论指数

2.1　量子化学基础知识

科学的进步和发展告诉我们,科学研究总是沿着从宏观到微观,从定性到定量,从观察、实验到理论化,从现象到本质的途径进行。当人们了解细胞结构之后,生物机体的生长和行为就可以得到一定程度的说明;当了解组成细胞的分子以及它们之间的反应之后,细胞本身就能被更好地认识了;当认识了组成分子的原子以及它们的电子结构和电子的运动规律之后,分子的性质就能更好地予以说明了。

通过第一章的讨论我们知道,要了解生命现象的实质,在电子水平上的研究是必不可少的。同时,我们也知道,发生在电子领域中的事件已经不能再用经典力学或经典化学的方法来描述。因为它的规律性是受量子力学所支配的,所以要用量子力学的方法来描述。但是,就量子生物学而论,运用量子化学方法更为直接。因为量子化学有着许多用电子结构分析分子结构和性质以及用量子理论计算电子结构的经验。

鉴于阐述量子力学基本原理的专著已有不少,且介绍量子化学计算方法的书籍也相继问世。为着使本书有一定的完整性,并着重考虑生物学、农学和医学院校或有关专业的读者,对量子理论和高等数学不太熟悉这样的实际,因此,在这一章里,我们仅对学习量子生物学所必须的量子力学(量子化学)基础知识和理论方法作扼要介绍。

2.1.1　波函数与轨道

2.1.1.1　波函数

微观粒子的运动不服从经典力学规律,而服从量子力学的规律。

量子力学的一个基本任务,就是求出微观粒子体系对应于一切可能状态的波函数 Ψ。确定波函数Ψ的微分方程,即Schrödinger方程,其一般形式为

$$\nabla^2\Psi+\frac{8\pi^2 m}{h^2}(E-V)\Psi=0 \tag{2.1}$$

也可写为

$$-\frac{h^2}{8\pi^2 m}\nabla^2\Psi+V\Psi=E\Psi \tag{2.2}$$

式(2.1)和(2.2)用来描述微观粒子运动的稳定状态,适用于微观粒子在空间出现的几率不随时间改变的情况。这对于讨论原子(或分子)稳定状态的核外电子运动是适用的,因为电子总是在

核外运动着，在稳定态时，电子在空间各处出现的几率不会随时间改变。所以Ψ也称为稳定状态的波函数。

Schrödinger方程也有用算符表示的。令

$$H=-\frac{h^2}{8\pi^2 m}\nabla^2+V \tag{2.3}$$

则得最简单形式：

$$H\Psi=E\Psi \tag{2.4}$$

式中H是一个称为Hamilton算符（或能量算符）的缩写符号。算符是一种运算符号（一个数学算符规定了一个或一组数学运算，如乘、除和微分）。

(2.4)式表示H作用到波函数Ψ上后就等于粒子的总能量与Ψ的乘积。E是这个微观粒子的总能量，是与Ψ相关的H的本征值。

波函数Ψ是描述微观粒子运动状态的函数，它包含着所描述体系的所能知道的全部知识。对于体系的一个特定状态，只要波函数Ψ已知，原则上就可以运用下述规则确定任何可观察物理量[1]：

$$可观察量=\frac{\int\Psi^*(算符)\Psi\,\mathrm{d}\tau}{\int\Psi^*\Psi\,\mathrm{d}\tau} \tag{2.5}$$

若可观察量为能量，就使用H，则(2.5)式可改写为

$$E=\frac{\int\Psi^*H\Psi\,\mathrm{d}\tau}{\int\Psi^*\Psi\,\mathrm{d}\tau} \tag{2.6}$$

若可观察量为电荷密度、偶极矩等，则用其它算符。

必须说明的是：Ψ只是一个普通的数学函数，是一个多自变量的函数。它的自变量就是这个微观粒子体系中各个粒子的位置坐标和自旋坐标。Ψ与它的共轭复数Ψ^*的乘积等于各个微观粒子在其坐标值规定的位置附近出现的几率密度，即$\Psi^*\Psi\,\mathrm{d}\tau$是在空间$\mathrm{d}\tau$体积元中找到粒子的几率，或$|\Psi|^2\mathrm{d}\tau$是体积元$\mathrm{d}\tau$内粒子出现的几率。

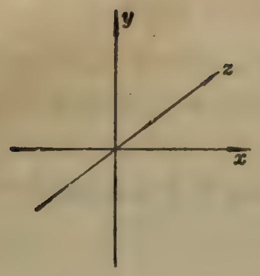
在量子力学中，因为涉及三维分子体系的问题，所以Ψ也随着这些坐标而变化。通常采用球坐标系来描述。球坐标系（r、θ和φ）与直角坐标系之间的变换关系如下（图 2-1）：

$$\begin{cases} x=r\sin\theta\cos\varphi \\ y=r\sin\theta\sin\varphi \\ z=r\cos\theta \end{cases}$$

$$r=\sqrt{x^2+y^2+z^2}$$

2.1.1.2 轨道

在量子生物学中，离不开应用原子轨道和分子轨道的概念，但我们最感兴趣的是后者。



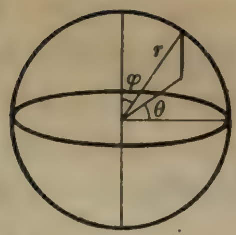


图 2-1 A. 直角坐标x、y和z，B. 球坐标r、θ和φ

所谓轨道，即是指单电子波函数。每个轨道都是一个描写单个电子行为的三维数学函数。

以一个原子核（或原子实）为中心的单电子波函数称为原子轨道 Φ，它是一个电子的空间坐标（球坐标r、θ和φ）的函数。换句话说，所谓原子轨道就是这样一些单电子波函数，它的绝对值平方表明了电子的几率密度分布。

所谓分子轨道，即是围绕着分子骨架分布的多中心单电子波函数。它同样也是单个电子的空间坐标的函数。分子轨道的概念是原子轨道的自然引伸，它假设把分子中的电子轨道扩展到整个分子。其实，分子轨道除开电子的运动所涉及的范围遍及分子中所有的原子外，它与原子轨道并无本质上的区别。然而，无论是原子轨道还是分子轨道，都只能在一定近似的基础上应用。

当我们假定两个或多个原子相互作用形成分子时，原来在各原子轨道上运动的电子即转入到整个分子的分子轨道上运动。原来是在单核多电子体系中运动的电子，现在转入到多核多电子体系中运动，而原来属于个别原子的电子，现在则属于整个分子。

在实际处理分子轨道时，我们可以用分子中各原子的原子轨道的线性组合来逼近分子轨道，这就是在量子化学中常用的原子轨道线性组合（LCAO）近似，其表示式为

$$\Psi = C_1\Phi_1 + C_2\Phi_2 + \cdots + C_n\Phi_n \qquad (2.7)$$

或

$$\Psi_i = \sum_{\mu} C_{\mu i}\Phi_{\mu} \qquad (\mu = 1, 2, 3, \cdots, n) \qquad (2.8)$$

如果分子轨道是以Φ_1和Φ_2的贡献为特征的，那么以原子轨道的线性组合（LCAO 近似）代表分子轨道的假设是合理的。于是，原子轨道线性组合方程为

$$\Psi = C_1\Phi_1 + C_2\Phi_2$$

其中，系数C_1和C_2是衡量二个原子轨道的贡献。

由于原子轨道线性组合是合理的，因此可以认为分子轨道是由原子轨道 Φ_1 和Φ_2 加和而得，且与 Φ_1 和 Φ_2结合的能量大致属于同一数量级。

2.1.2 量子化学中的基本近似

迄今为止的事实说明，对多于二粒子的系统，Schrödinger 方程的严格解是非常困难的。所有分子波函数都是近似的，只是近似程度不同而已。在量子化学中作了三个近似处理，即引入了非相对论近似、Born-Oppenheimer 近似和轨道近似。

由于量子生物学处理的分子——生物分子和药物分子，是复杂的多粒子体系，因而必须采取各种近似。

2.1.2.1 非相对论近似

电子在原子核附近运动但又不被原子核俘获，必须保持很高的运动速度。依据相对论，此时电子的质量μ不是一个常数，而由电子运动速度 v 、光速 c 和电子静止质量μ_0决定

$$\mu=\frac{\mu_0}{\sqrt{1-\frac{v_2}{c}}} \tag{2.9}$$

非相对论近似忽略这一相对论效应，认为电子质量$\mu=\mu_0$。

非相对论量子力学为研究微观粒子体系提供了运动方程，对于我们所感兴趣的分子的电子结构，则需求解一个与时间无关的Schrödinger方程(2.4)式。这是一个多粒子体系问题，其中的Ψ是与分子中所有电子和原子核的坐标有关的一个可能稳定态的波函数。Hamilton 算符$\boldsymbol{H}$则包括电子的动能和排斥能(采用原子单位制，以下同)，原子核的动能，电子与核的吸引以及核的排斥等项

$$\boldsymbol{H}=-\frac{1}{2}\sum_p\frac{1}{M_p}\nabla^2-\frac{1}{2}\sum_i\nabla_i^2+\frac{1}{2}\sum_{i\neq j}\frac{1}{r_{ij}}$$
$$-\sum_p\sum_i\frac{Z_p}{r_{pi}}+\frac{1}{2}\sum_{p\neq q}\frac{Z_pZ_q}{R_{pq}} \tag{2.10}$$

式中，指标i、j指电子，指标p、q指核，Z、M分别表示核的电荷及质量，r、R均表示距离。

(2.10)式所表示的Hamilton 不仅描写电子绕核的运动，而且也描述原子核的相对运动以及分子作为一个整体在空间的运动。将(2.10)式代入(2.4)式就得到用原子单位表示的Schrödinger方程

$$\left\{-\frac{1}{2}\sum_p\frac{1}{M_p}\nabla^2-\frac{1}{2}\sum_i\nabla_i^2+\frac{1}{2}\sum_{i\neq j}\frac{1}{r_{ij}}\right.$$
$$\left.-\sum_p\sum_i\frac{Z_p}{r_{pi}}+\frac{1}{2}\sum_{p\neq q}\frac{Z_pZ_q}{R_{pq}}\right\}\Psi=E\Psi \tag{2.11}$$

在(2.11)式中，把电子的质量视为其静止质量μ_0，这仅在非相对论条件下才成立，称为非相对论近似。

2.1.2.2 Born-Oppenheimer近似

由于组成分子体系的原子核的质量比电子的质量大10^3 至10^5 倍，而电子的运动速度又比核

的运动速度大得多，这就允许把核的运动和电子的运动分离开来，即在求解电子问题时，近似地认为原子核是固定不动的，这种固定核的近似就称为Born-Oppenheimer近似。

对于固定的原子核，(2.10)式中第五项核间排斥势是常数。假定分子波函数可以表示为电子运动波函数和核运动波函数的乘积，亦即 $\Psi_e(\boldsymbol{r},\boldsymbol{R})\Psi_n(\boldsymbol{R})$，则方程(2.4)或(2.11)通过分离变量而得电子运动的 Schrödinger 方程

$$\left\{-\frac{1}{2}\sum_i\nabla_i^2+\frac{1}{2}\sum_{i\neq j}\frac{1}{r_{ij}}-\sum_p\sum_i\frac{Z_p}{r_{pi}}\right\}\Psi_e(\boldsymbol{r},\boldsymbol{R})=\boldsymbol{E}(\boldsymbol{R})\Psi_e(\boldsymbol{r},\boldsymbol{R}) \tag{2.12}$$

电子运动的总能量 $\boldsymbol{E}(\boldsymbol{R})$是作为核运动方程的势能。以后，我们将只讨论方程(2.12)，仍记作

$$\boldsymbol{H}\Psi=\boldsymbol{E}\Psi$$

或

$$\boldsymbol{H}\Psi_e(\boldsymbol{r},\boldsymbol{R})=\boldsymbol{E}(\boldsymbol{R})\Psi_e(\boldsymbol{r},\boldsymbol{R}) \tag{2.13}$$

其中

$$\boldsymbol{H}=-\frac{1}{2}\sum_i\nabla_i^2+\frac{1}{2}\sum_{i\neq j}\frac{1}{r_{ij}}-\sum_p\sum_i\frac{Z_p}{r_{pi}} \tag{2.14}$$

2.1.2.3 轨道近似

对于多电子体系，按照式(2.14)表示的 $\boldsymbol{H}$，方程(2.13)仍不可能严格求解。问题在于(2.14)式中多电子 Hamilton 算符 包含了$\frac{1}{r_{ij}}$形式的电子相互作用项，故不能分离变量。

当电子间相互作用不存在时，(2.14)式可约化为单电子算符的和

$$\boldsymbol{H}_0=\sum_{i=1}^{N}h(\boldsymbol{r}_i) \tag{2.15}$$

其中

$$h(\boldsymbol{r}_i)=-\frac{1}{2}\nabla_i^2-\sum_p\frac{Z_p}{r_{pi}} \tag{2.16}$$

为有固定位置的核存在时，一个单电子的 Hamilton_0 $\boldsymbol{H}_0$ 的本征函数是各轨道(单电子波函数)的乘积

$$\psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2)\cdots\psi_N(\boldsymbol{r}_N) \tag{2.17}$$

而各轨道即为单电子算符 h 的本征函数

$$h(\boldsymbol{r})\psi_n(\boldsymbol{r})=\epsilon_n\psi_n(\boldsymbol{r}) \tag{2.18}$$

在自旋和反对称性加入以后，轨道乘积式(2.17)就是该体系某些状态时的简单的轨道近似。一般地，轨道近似是一个 N 电子体系的模型，其中各电子的运动（或状态）是用仅为该电子坐标的函数的自旋-轨道来描写的。通常，这种自旋-轨道具有形式

$$\psi_n'(\boldsymbol{r},\sigma)=\psi_n(\boldsymbol{r})\eta_n(\sigma) \tag{2.19}$$

(2.19)式中 ψ_n 描述电子的空间运动，自旋函数 η_n 描写电子的自旋 状 态。于是，对于一个 N 电子体系的总波函数，就是 N 个自旋-轨道的反对称的乘积

$$\Psi(1,2,\cdots,N)=MA_N[\psi'_1(1)\psi'_2(2)\cdots\psi'_n(N)] \tag{2.20}$$

式中 $\psi'_n(i)=\psi_n(i)\eta_n(i)$是电子 i 的坐标的函数，M 是归一化常数，A_N 是反对称算符。这个波函数是相应于在N个自旋-轨道中有 $N!$ 种分配电子方式的 $N!$ 个自旋-轨道乘积的线性组合，它可以写成 Slater 行列式的形式

$$\Psi(1,2,\cdots,N)=\frac{1}{\sqrt{N!}}\begin{vmatrix}\psi'_1(1)\psi'_1(2)\cdots\psi'_1(N)\\ \psi'_2(1)\psi'_2(2)\cdots\psi'_2(N)\\ \vdots \quad \vdots \quad \vdots\\ \psi'_N(1)\psi'_N(2)\cdots\psi'_N(N)\end{vmatrix} \tag{2.21}$$

简记为

$$\Psi(1,2,\cdots,N)=\frac{1}{\sqrt{N!}}D\left|\psi'_1(1)\psi'_2(2)\cdots\psi'_N(N)\right| \tag{2.22}$$

其中$\frac{1}{\sqrt{N!}}$为归一化系数(这里假定每个单电子波函数都是归一化的)。

一组N个自旋-轨道可以唯一地确定(2.21)式的波函数，但反过来，同一个波函数 Ψ 却不止由一组自旋-轨道所确定，这是由行列式的某些性质所决定的。为此，我们可以构成无限多个可供选择的组，其中必有一个集合的自旋-轨道常常可以选成正交归一化的：

$$\int\psi'^*_m\psi'_n\mathrm{d}\tau=\delta_{mn} \tag{2.23}$$

在轨道近似下讨论多电子体系的问题，常用到原子轨道线性组合成分子轨道和自洽场的方法(见后面)。

实践证明，对许多问题，只要选择适当的原子轨道基函数，采用单行列式波函数就能得到对体系准确波函数的简单而又比较满意的近似。

2.2 量子生物学方法和指数

在这一节里，我们仅限于讨论量子生物学的理论计算方法和指数。

大量的实验和事实表明，生物分子所产生的生物学作用跟分子本身的物理、化学性质是密切相关的，而且生物分子的特异生物活性在一定程度上又是生物分子电子结构的函数。所谓电子结构，通常包括分子的电子分布，立体化学，静态和动态的能量分布。对于这些结构特征的确定和评价正是量子生物学家探索生物分子活性受电子结构影响的中心课题。

在量子生物学中确定和评价各个分子的电子结构和反应性能时，根据研究对象和目的，结合我们的计算条件，首先就是要选择量子化学计算方法和确定理论指数，下面我们将分别介绍常用的计算方法和计算指数。

2.2.1 简单分子轨道法

简单分子轨道法的特点是不明显地考虑电子间的相互作用。

2.2.1.1 Hückel 分子轨道法(HMO 法)

HMO 法是由 Hückel 在 1931 年提出的。这是一个经验的处理 π 电子体系 的 理 论 近似方法,主要用于 π 电子体系,即平面共轭分子体系(如核酸的嘌呤和嘧啶碱基、芳香族氨基酸等即是具有共轭结构的生物分子)。

1. 计算原理[2]

在分子轨道理论中,对一个分子中参加共轭的原子系统,根据原子轨道线性组合及线性变分法,可以建立 Hückel 矩阵(H):

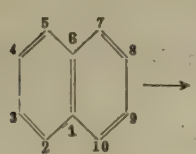
$$(H)=\begin{pmatrix} H_{11} & H_{12}\cdots & H_{1n} \\ H_{21} & H_{22}\cdots & H_{2n} \\ \vdots & \vdots & \vdots \\ H_{n1} & H_{n2}\cdots & H_{nn} \end{pmatrix} \tag{2.24}$$

按照 Hückel 理论的假设,它的矩阵元分别为

$$(H_{ij})\begin{cases} \left.\begin{matrix} 0 & \alpha_C=0 \\ \dfrac{\alpha_i}{\beta_{C-C}}=\delta_i & \delta_i\neq 0 \end{matrix}\right\} i=j \text{ 库仑项} \\ \text{不相邻原子均为 } 0 \\ \left.\text{相邻原子}\begin{cases} \dfrac{\beta_{C-C}}{\beta_{C-C}}=1 \\ \dfrac{\beta_{ij}}{\beta_{C-C}}=\eta_{ij} \end{cases}\right\} i\neq j \text{ 共振项} \end{cases} \tag{2.25}$$

其中 $\alpha_i=\int\Phi_i H\Phi_i \mathrm{d}\tau$ 为第 i 个原子的库仑积分,取碳原子 $\alpha_C=0$,$\beta_{ij}=\int\Phi_i H\Phi_j \mathrm{d}\tau$ 为第 i 个原子和第 j 个原子的共振积分。在 HMO 理论中,令不相邻原子间共振积 分 为零。δ_i 为第 i 个杂原子库仑积分经验参数,η_{ij} 为第 i 个杂原子与第 j 个原子的共振积分经验参数。

实际上,在全碳原子的共轭体系(如苯、萘)中,Hückel 矩阵是一个 分 子 图 中的原子连接矩阵,亦称拓扑矩阵,用 0 和 1 表示不相连接与相连接,我们就可以容易地建立全碳原子共轭体系的 Hückel 矩阵,如萘(图 2-2):



	1	2	3	4	5	6	7	8	9	10
1	0	1	0	0	0	1	0	0	0	1
2	1	0	1	0	0	0	0	0	0	0
3	0	1	0	1	0	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0	0
5	0	0	0	1	0	1	0	0	0	0
6	1	0	0	0	1	0	1	0	0	0
7	0	0	0	0	0	1	0	1	0	0
8	0	0	0	0	0	0	1	0	1	0
9	0	0	0	0	0	0	0	1	0	1
10	1	0	0	0	0	0	0	0	1	0

图 2-2 萘的分子图

在图 2-2 中，第 i 行包含第 i 原子和其它原子连接的信息。如第一行即表示第 1 个碳原子分别与第 2、6、10 号碳原子相连。

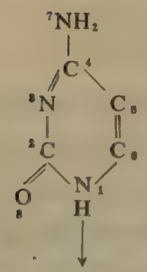
但是，如果在分子中包含有杂原子，则可将杂原子位置 i 的对角元用相应的 δ_i 替代 0，并将相应的共振积分参数 η_{ij} 替代 1。于是，同样可以容易地建立含杂原子分子的 Hückel 矩阵。在 HMO 法中，杂原子积分参数的选择见表 2-1[2]。

表 2-1 HMO 杂原子参数表

类号	原子	结合形式	库仑积分 δ	结合形式	共振积分 η
1	C	=C—	0	—C=C—	1
2	N	=N—	0.4	—C=N—	1
3		—N̈—	1	=C—N̈—	0.9
4		=N⁺—	2	—C=N⁺—	1
5	O	=O	1.2	—C=O	2
6		—Ö—	2	=C—Ö—	0.9
7				=N—Ö—	0.69
8	C^a—C^b—H_3*	=C^a—	−0.1	C^a—C^b	0.7
9		≡H_3	−0.2	—C≡H_3	2
10	S	=S	0	—C=S	1.2
11		—S̈—	0	=C—S̈—	0.6
12	—N(=O)(O)		$\delta_O=\delta_N=1$	—N⩵O	1
13	F	—F	3	=C—F	0.7
14	Cl	—Cl	2	=C—Cl	0.4
15	Br	—Br	1.5	=C—Br	0.3
16	P	—P—	−0.2	—P—O	0.7
17	B	C—B	−1.0	C—B	0.7

* C^a 为芳香族；C^b 为脂肪族

现以胞嘧啶碱基为例，按以上所述，我们有：



$$\begin{bmatrix} x+1 & 0.9 & 0 & 0 & 0 & 0.9 & 0 & 0 \\ 0.9 & x & 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & x+0.4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & x & 1 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 1 & x & 1 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 1 & x & 0 & 0 \\ 0 & 0 & 0 & 0.9 & 0 & 0 & x+1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & x+1.2 \end{bmatrix} = 0$$

有了 Hückel 矩阵，即可建立本征方程，

$$HC = C\lambda \tag{2.26}$$

其中 C 是本征向量组成的矩阵，它的每一列代表由原子轨道组成的分子轨道分量

$$\psi_i = \sum C_{i\nu}\Phi_\nu \tag{2.27}$$

λ 是本征值 λ_i 组成的对角阵。由 λ_i 可得分子轨道能级，它们可以方便地依据线性代数中求解本征值和本征向量的 JACOB 法或 QL 叠代法，通过计算机获得。由 C（即波函数）可得一系列量子化学指数（或量子生物学指数）。

2．理论指数和计算公式

(1) 电荷密度和净电荷

一个电子在微体积元 $d\tau$ 内出现的几率为 $|\Psi|^2 d\tau$。在原子轨道线性组合的 Hückel 近似中，对分子轨道中的每个原子轨道的波函数作同样的假定，它对分子轨道波函数 Ψ 的贡献由系数 C_i 来衡量。因此，电子在与原子轨道 Φ_i 有关的空间出现的几率是 C_i^2。在原子 r 附近，这个电子出现的几率为

$$q_{ir} = C_{ir}^2 \tag{2.28}$$

如果体系中有一系列的分子轨道波函数 $\psi_1, \psi_2, \psi_3, \cdots, \psi_k$，在相应的轨道上有 n_k 个电子，则原子 r 的电荷密度是

$$\begin{aligned} q_r &= n_1 C_{1r}^2 + n_2 C_{2r}^2 + \cdots + n_k C_{kr}^2 \\ &= \sum n_i C_{ir}^2 \end{aligned} \tag{2.29}$$

(2.29)式对于分子的基态、激发态、离子或游离基都是适用的。由于一个分子轨道最多能填充两个电子，因此 n_i 可以是 0 、1 或 2 。但在通常情况下分子在基态时占据轨道中 n_i 都是 2，故原子 r 的总的电荷密度是它对所有占据分子轨道的电荷密度之和

$$q_r = \sum_{i=1}^{OCC*} 2C_{ir}^2 \qquad (2.30)$$

但是，有时为了表示出分子中原子上实际的荷电情况，常用净电荷 Q_r 代替电荷密度 q_r

$$Q_r = K_r - q_r \qquad (2.31)$$

(2.31)式中，K_r 为第 r 个原子所提供的 π 电子数，如果 r 原子能提供孤对电子，则 $K_r = 2$。

除电荷密度和净电荷指数外，还有所谓前沿电荷密度：

$$f_r^E = n_i C^2_{HOMOr} \qquad (2.32)$$

$$f_r^N = n_i C^2_{LUMOr} \qquad (2.33)$$

式中，f^E、f^N 分别表示亲电、亲核前沿电荷密度。足标 HOMO、LUMO 则分别表示最高占据分子轨道(highest occupied molecular orbital) 和最低空轨道 (lowest unoccupied molecular orbital)。

(2) 最高占据轨道(HOMO)和最低空轨道(LUMO)的能级。

HOMO 和 LUMO 在讨论分子的反应性能时有着特殊的意义。当分子间以形成电荷迁移络合物的方式相互作用时，HOMO 能级可以作为分子给电子能力的量度，亦即电子是从供体的 HOMO 迁移到受体的 LUMO。

在 HMO 法中能量是以 β（负值）为单位的。因此，HOMO 的值越大，则表示这一轨道中的电子越稳定，该分子给电子的能力越小。而 LUMO 值越大，则表示电子进入该轨道后体系能量降低得越多，即该分子接受电子的能力越强。

我们知道，用实验方法确定生物学和药学中有意义的化合物的电离势和电子亲合势是比较困难的。但是，HOMO 和 LUMO 的能级却与电离势和电子亲合势有着密切关系。因此，可以通过计算分子的 HOMO 和 LUMO 能级来近似地了解分子的电离势和电子亲合势，这在药物分子等的设计中是十分重要的。

关于 HOMO 和 LUMO 能级的计算，读者可参见计算实例。

(3) 超离域度

超离域度概念包括近似超离域度(S_r')、亲电超离域度(S_r^E)、亲核超离域度(S_r^N)以及自由基反应的超离域度。其中：

$$S_r^E = 2\sum_i^{OCC} \frac{(C_{ir})^2}{\lambda_i} (\varepsilon_i = \alpha + \lambda_i \beta) \qquad (2.34)$$

$$S_r^N = 2\sum_i^{VAC} \frac{(C_{ir})^2}{-\lambda_i} \qquad (2.35)$$

$$S_r^R = \sum_i^{OCC} \frac{(C_{ir})^2}{\lambda_i} + \sum_i^{VAC} \frac{(C_{ir})^2}{-\lambda_i} \qquad (2.36)$$

在(2.34)至(2.36)式中，C_{ir}是第 i 个分子轨道中第 r 个原子轨道的系数。

实践表明，超离域度不仅是分子内反应难易程度的尺度，同时也是比较在分子之间反应性大

* OCC 为占据轨道，VAC 为空轨道。

小时的一种理论指数。由于它与很多实验结果相对应，从而扩大了人们对于生物分子的反应活性以及对各类药物作用的认识。在作者等[3]的工作中，以超离域度为判据探讨多环芳烃（PAH）及其甲基取代衍生物的电子结构与致癌活性的关系，取得了与实验致癌活性相符的结果。但在运用这一指数时要注意到它的某些不足和可能出现的问题。

（4）键级

如果分子中，某一电子所在的分子轨道为Ψ_i：

$$\Psi_i = C_{i1}\Phi_1 + C_{i2}\Phi_2 + \cdots + C_{in}\Phi_n$$

则乘积 $C_{i1}C_{i2}$ 可以作为电子云分布在 1 、2 两个原子之间的尺度，也就是这个电子贡献给原子 1 和 2 的一部分键级。

在 Ψ_i 分子轨道中，成键原子 r 和 s 之间的部分键级是

$$P^i_{rs} = C_{ir}C_{is} \tag{2.37}$$

而两原子间的总键级（简称键级）是所有占据分子轨道贡献之和，故有

$$P_{rs} = \sum_i^{\text{occ}} n_i C_{ir} C_{is} \tag{2.38}$$

对于充满的分子轨道，$n_i = 2$。

键级的大小，可以认为是和一个键的成键能力相关联的。键级的数值越大，键的强度亦大，则键长越短。当所有的系数都较大，而且符号相同，则乘积为正，表示两邻接原子间有较强的成键能力。如果两个系数中只有一个值比较大，则说明电子主要位于一个核的周围，而不是在两个核之间。如果两个系数中有一个系数为零，则部分键级为零。于是，在该原子本身所处的地方就是一个节点，说明电子对于成键没有贡献。当两原子间系数的符号不相同时，则部分键级为负，两原子之间有一个节点，这就是反键的。

键级的一个重要应用就是预测键长。键级的数值越大键长则越短。对简单分子，碳-碳键级与实验键长有良好线性关系，但是对于生物分子和其它较为复杂的分子，这种线性关系并不存在。

此外，我们也可以运用下列经验公式来估算键长：

$$R(\text{nm}) = 0.150 - 0.016\,p \tag{2.39}$$

式中，p 为 π 键级。

Coulson 和 Longuet-Higgins[4]还提出，π 键键能可以用电荷密度和键级的函数来表示：

$$E_\pi = \sum_r q_r \alpha + 2\sum_{r<s}\sum p_{rs}\beta \tag{2.40}$$

这个公式在校正计算值时很有用，因为求得系数 C_r 以后很容易求得 q_r 和 p_{rs}，将q_r 和 p_{rs} 代入(2.40)式即可求得总的 π 键键能。

(2.40)式表明了共轭分子中各个键的键级加和对于总能量的关系；如果一个原子与其邻近原子间的键级之和大，对于总能量的贡献亦大，可以设想该原子在分子中将受到较强的制约而不

易反应，表现出较低的反应活性。反之，则活性较高。

(5) 自由价

“自由价”是 Coulson[5]定义的一个反应活性指数。它是指原子的 π 电子的残余键合力，是原子最大成键度 N_{max} 和实际成键度 $\sum_s p_{rs}$ 之差：

$$F_r = N_{max} - \sum_s p_{rs} \tag{2.41}$$

式中 s 是与原子 r 相邻接的原子；N_{max} 系经验值。按文献所列并考虑从 Pauling 电负性出发，将硫、磷和卤素原子分别处理为

$$S_{N_{max}} \doteq P_{N_{max}} \doteq Br_{N_{max}} \doteq C_{N_{max}};$$

$$Cl_{N_{max}} \doteq N_{N_{max}};\quad F_{N_{max}} \doteq O_{N_{max}}$$

在实际计算时，对于不同的原子，N_{max} 可分别取作：

$$N_{max} = \begin{cases} \sqrt{3} & (\text{对碳原子}) \\ \sqrt{2} & (\text{对氮原子}) \\ 1 & (\text{对氧原子}) \end{cases} \tag{2.42}$$

但是，“自由价”的概念一般只限于用在碳原子，对于氮、氧和氯等含有未共用电子对的原子，自由价的概念就不那么明确，还有待作进一步探讨。

在一般情况下，对于我们计算的分子，F_r 值越大的碳原子，其残余键合能力也越大，预期它的反应活性也将会越高。

3. 计算实例

现在以核酸的组分碱基之一——胞嘧啶(cytosine)为例，简略介绍 HMO法基础上的理论指数计算方法。依照 2.2.1.1，我们很容易从胞嘧啶的结构式建立 Hückel 矩阵，进而运用JACOB法(或其它求解本征值和本征向量的方法)在计算机上算出本征值和本征向量。最后，按照前面所述的理论指数计算公式，可得胞嘧啶的若干理论指数。

用 HMO 法计算得到的胞嘧啶的本征值和本征向量如表 2-2 所示。

表 2-2 胞嘧啶(内酰胺式)的本征值和本征向量*

y	−3.0682	−2.0077	−1.5920	−0.7847	−0.5951
C_1	0.3146	0.1129	0.6810	0.2669	−0.4158
C_2	0.6037	−0.1269	−0.0454	−0.0273	−0.1206
C_3	0.2765	0.2714	−0.2215	0.5243	0.4950
C_4	0.1342	0.5631	−0.2186	0.1744	0.0240
C_5	0.0826	0.4005	0.1726	0.2689	0.5573
C_6	0.1192	0.2530	0.4934	0.0366	0.3076
C_7	0.0584	0.5029	−0.3323	−0.7292	−0.0534
C_8	0.6463	−0.3137	−0.2318	−0.1314	0.3987

* 本表仅列出占据分子轨道能级和系数。

由表 2-2 和理论指数计算公式可计算：

（1）电荷密度（以N_3为例）

$$q_{N_3}=2(0.2765)^2+2(0.2714)^2+2(-0.2215)^2+2(0.5243)^2+2(0.4950)^2=1.4382$$

净电荷

$$Q_{N_3}=1-1.4382=-0.4382$$

（2）键级（以N_1—C_2键为例）

$$p_{12}=2(0.3146)(0.6037)+2(0.1129)(-0.1267)+2(0.6810)(-0.0454)+2(0.2669)(-0.0273)+2(-0.4158)(-0.1206)=0.3751$$

（3）自由价（以N_1为例）

$$F_{N_1}=\sqrt{2}-(p_{16}+p_{12})=\sqrt{2}-(0.529+0.3751)=0.510$$

（4）亲电超离域度（以N_1位置为例）

$$S^E_{N_1}=2\left[\frac{(0.3146)^2}{-3.0682}+\frac{(0.1129)^2}{-2.0077}+\frac{(0.6810)^2}{-1.5920}+\frac{(0.2669)^2}{-0.7847}+\frac{(-0.4158)^2}{-0.5951}\right]=-1.4224$$

（5）最高占据轨道和最低空轨道能级

$$\text{HOMO}=-0.5951$$
$$\text{LUMO}=-0.7949$$

通过理论计算概括起来可以得到两类指数，一类是能量指数，一类是结构指数。能量指数可说明体系的能量状态。结构指数可说明体系中各原子所在处的电荷，结合能力以及键的性质，从而用以讨论生物分子的性质与行为。如能量指数中的HOMO反映电离势，而LUMO 反映电子亲合势，用以说明体系得失电子的能力。结构指数中的键级、自由价、原子净电荷等又都与化学反应能力有关。

研究的实践表明，以上例举的一些理论指数（参见2.2.1.1,2），在说明和预言生物分子或药物分子的电子结构和反应活性间的关系时，是有用的。特别作为理论计算指数它们能够提供由实验不可能得到的分子内部电荷分布的信息。在量子生物学中，如果我们对分子体系之间的差别更感兴趣，那么理论指数会更有意义。但是，这些理论指数都是从未受微扰分子的计算中推导出来的。因此，在运用时应慎重行事，必须非常仔细地检验它们作为反应活性参数的若干性

质，以求与实验结果相一致。尤其是由对孤立分子的计算而得到的理论指数都是静态的指数，在这种情况下，它们对于说明动态的生物分子的反应活性显然存在着潜在的危险性。再者，热力学的研究告诉我们，反应活性依赖于过渡态而不依赖于反应物。这就使我们更加认识到上述理论指数应是有条件的，它还给我们留下了不少尚待解决的问题。

必须指出，前面已经介绍过的理论指数，最初都是在HMO 方法的基础上提出来的，现在大都已被推广到其它更精确的计算方法当中。因此，在以下讨论其它理论方法时，在一般情况下，我们将不再重复介绍这些理论指数。

2.2.1.2 微扰分子轨道法(PMO法)[6,7]

量子化学在处理较大分子时，采用微扰法求解Schrödinger方程。

理论计算的研究表明，我们在这里讨论的微扰分子轨道法(简称PMO法)，对于分子稳定性和反应的计算在实用上是一个简单而有力的工具。这个方法所作近似的目的，是直接计算体系间的差异，而不是将体系分别处理后再确定其差异。微扰理论可以研究两个体系间的静电作用、电荷转移和Van der Waals 力等相互作用。

1. 计算原理

在这一节里，我们仅介绍定态非简并微扰(对于有简并时的微扰，读者可从有关的量子化学著作中了解到)。

在量子力学中，微扰理论的出发点是 Schrödinger 方程式，即当体系受到微扰作用时，其Hamilton算符$\boldsymbol{H}$可分解为

$$H=H_0+\lambda H' \tag{2.43}$$

(2.43)式中$\lambda H'$为微扰项。于是，该体系的Schrödinger
方程可写成

$$(H_0+\lambda H')\psi_n=E_n\psi_n \tag{2.44}$$

式中，ψ_n、E_n为微扰体系的波函数和本征值。按照Taylor级数展开的方式，在确定这些级数展开是收敛的前提下，可将E_n和ψ_n对λ的幂级数展开，即：

$$\psi_n=\psi_n^{(0)}+\lambda\psi_n^{(1)}+\lambda^2\psi_n^{(2)}+\cdots\cdots \tag{2.45}$$

$$E_n=E_n^{(0)}+\lambda E_n^{(1)}+\lambda^2E_n^{(2)}+\cdots\cdots \tag{2.46}$$

将(2.45)和(2.46)代入(2.44)式得：

$$\begin{aligned}&H_0\psi_n^{(0)}+\lambda(H'\psi_n^{(0)}+H_0\psi_n^{(1)})+\lambda^2(H'\psi_n^{(1)}+H_0\psi_n^{(2)})+\cdots\cdots\\=&E_n^{(0)}\psi_n^{(0)}+\lambda(E_n^{(1)}\psi_n^{(0)}+E_n^{(0)}\psi_n^{(1)})+\lambda^2(E_n^{(2)}\psi_n^{(0)}+E_n^{(1)}\psi_n^{(1)}\\&+E_n^{(0)}\psi_n^{(2)}+\cdots\cdots\end{aligned} \tag{2.47}$$

此方程要使λ在所有数值变化中的上述等式成立，就要求等式两边有相同幂次方的λ系数相等，即：

$\lambda=0$项

$$H_0\psi_n^{(0)}=E_n^{(0)}\psi_n^{(0)} \tag{2.48}$$

λ'的项

$$(H_0-E_n^{(0)})\psi_n^{(1)}=E_n^{(1)}\psi_n^{(0)}-H'\psi_n^{(0)} \tag{2.49}$$

λ^2的项

$$(H_0-E_n^{(0)})\psi_n^{(2)}=E_n^{(2)}\psi_n^{(0)}+E_n^{(1)}\psi_n^{(1)}-H'\psi_n^{(1)} \tag{2.50}$$

(2.48)式即为无微扰体系的 Schrödinger 方程，求解后就可得到零级近似的能级 $E_n^{(0)}$和波函数 $\psi_n^{(0)}$，再从(2.49)式解出一级近似的能量校正值$E_n^{(1)}$和波函数$\psi_n^{(1)}$，由（2.50）式又可解出二级近似的校正值$E_n^{(2)}$和$\psi_n^{(2)}$等等。

由(2.48)式求得一级微扰体系的能量和波函数：

$$E_n^{(1)}=\int\psi_n^{*(0)}H'\psi_n^{(0)}\mathrm{d}\tau=H'_{nn} \tag{2.51}$$

$$\begin{aligned}\psi_n^{(1)}&=\sum_{m\neq n}\frac{\int\psi_m^{*(0)}H'\psi_n^{(0)}\mathrm{d}\tau}{E_n^{(0)}-E_m^{(0)}}\psi_m^{(0)}\\&=\sum_{m\neq n}\frac{H'_{mn}}{E_n^{(0)}-E_m^{(0)}}\psi_m^{(0)}\end{aligned} \tag{2.52}$$

将(2.51)和(2.52)式代入(2.50)式，可求得二级微扰项。在量子化学（量子生物学）中，最常使用的是一级和二级微扰项。其中包括二级微扰的本征值和波函数为

$$E_n=E_n^{(0)}+H'_{nn}+\sum\frac{H'_{nm}H'_{mn}}{E_n^{(0)}-E_m^{(0)}}+\cdots\cdots \tag{2.53}$$

$$\begin{aligned}\psi_n=\psi_n^{(0)}&+\sum_{m\neq n}\frac{H'_{mn}}{E_n^{(0)}-E_m^{(0)}}\psi_m^{(0)}+\sum_{k\neq n}\Bigg[\sum_{m\neq n}\frac{H'_{km}H'_{mn}}{(E_n^{(0)}-E_k^{(0)})(E_n^{(0)}-E_m^{(0)})}\\&-\frac{H'_{nn}H'_{kn}}{(E_n^{(0)}-E_k^{(0)})^2}\Bigg]\psi_k^{(0)}-\frac{1}{2}\sum_{m\neq n}\left(\frac{H_{mn}}{E_n^{(0)}-E_m^{(0)}}\right)^2\psi_n^{(0)}+\cdots\cdots\end{aligned} \tag{2.54}$$

在HMO法中，当共轭的R体系受到微扰后，则其第r个原子的库仑积分α_r改变为$\alpha_r+\delta\alpha_r$，$\delta\alpha_r=\int\Phi_rH'\Phi_r\mathrm{d}\tau$，由(2.53)、(2.54)式经整理后可得闭壳层共轭体系总电子能量和电荷密度的改变为*

$$\begin{aligned}\delta E_a&=2\sum_i^{\text{OCC}}(C_{ir})^2\delta\alpha_r\\&+2\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}}\frac{(C_{ir}C_{kr})^2}{\varepsilon_i-\varepsilon_k}(\delta\alpha_r)^2+\cdots\cdots\end{aligned} \tag{2.55}$$

$$\delta q_a=4\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}}\frac{C_{ir}C_{is}C_{kr}C_{ks}}{\varepsilon_i-\varepsilon_k}\delta\alpha_r+\cdots\cdots \tag{2.56}$$

同理，体系受到微扰后，其共振积分β_{rs}改变为$\beta_{rs}+\delta\beta_{rs}$，$\delta\beta_{rs}=\int\Phi_rH'\Phi_s\mathrm{d}\tau$，则电子总能量改变为

---

* 按照不相容原则，每个分子轨道最多只能容纳两个电子且自旋必须相反。若分子中的壳层不是全满就是全空，则称为闭壳层分子，对应的组态称为闭壳层组态；否则称为开壳层分子和开壳层组态。

$$\delta E_\beta = 2p_{rs}\delta\beta_{rs} + 2\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{(C_{ir}C_{ks}+C_{is}C_{kr})^2}{\varepsilon_i-\varepsilon_k}(\delta\beta_{rs})^2+\cdots\cdots \tag{2.57}$$

在(2.55)至(2.57)中，UNOCC为未占据轨道数，在(2.57)式中p_{rs}为r原子与s原子间的键级。而电荷密度改变为

$$\delta q_\beta = 4\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{(C_{ir}C_{ks}+C_{is}C_{kr})C_{it}C_{kt}}{\varepsilon_i-\varepsilon_k}\delta\beta_{rs}+\cdots\cdots \tag{2.58}$$

为使计算简洁，兹定义：

原子-原子极化率

$$\pi_{rs} = \frac{\partial q_r}{\partial \alpha_s} = 4\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{C_{ir}C_{is}C_{kr}C_{ks}}{\varepsilon_i-\varepsilon_k}$$

原子自极化率

$$\pi_{rr} = \frac{\partial q_r}{\partial \alpha_r} = 4\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{(C_{ir}C_{kr})^2}{\varepsilon_i-\varepsilon_k}$$

原子-键极化率，键-原子极化率

$$\pi_{r,st} = \frac{\partial q_r}{\partial \beta_{st}} = \frac{\partial^2 E}{\partial \alpha_r \partial \beta_{st}}$$

$$\pi_{st,r} = \frac{\partial p_{st}}{\partial \alpha_r} = 2\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{(C_{is}C_{kt}+C_{it}C_{ks})C_{is}C_{kr}}{\varepsilon_i-\varepsilon_k}$$

且可证明$\pi_{r,st}=2\pi_{st,r}$

键-键极化率

$$\pi_{rs,tu} = 2\sum_i^{\text{OCC}}\sum_k^{\text{UNOCC}} \frac{(C_{ir}C_{ks}+C_{is}C_{kr})(C_{it}C_{ku}+C_{tu}C_{kt})}{\varepsilon_i-\varepsilon_k}$$

2. 理论指数和计算公式

由上述定义，我们可以获得R共轭体系中某r原子轨道α_r发生了$\delta\alpha_r$变化时，能量E、电荷q和键级p的变化：

$$\delta E \doteq q_r\delta\alpha_r + \frac{1}{2}\pi_{rr}(\delta\alpha_r)^2 \tag{2.59}$$

$$\delta q \doteq \pi_{r,s}\delta\alpha_r \tag{2.60}$$

$$\delta p_{st,r} \doteq \pi_{st,r}\delta\alpha_r = \frac{1}{2}\pi_{r,st}\delta\alpha_r \tag{2.61}$$

当键$r-s$的β_{rs}发生了$\delta\beta_{rs}$变化时，能量E、电荷q和键级p的变化为

$$\delta E \doteq 2p_{rs}\delta\beta_{rs} + \pi_{rs,rs}(\delta\beta_{rs})^2 \tag{2.62}$$

$$\delta q \doteq \pi_{r,st}\delta\beta_{rs} = 2\pi_{st,r}\delta\beta_{rs} \tag{2.63}$$

$$\delta p_{tu} \doteq \pi_{rs,tu}\delta\beta_{rs} \tag{2.64}$$

上述六个量子化学理论指数即为PMO法获得的结果。从(2.59)至(2.64)式不难看出，所有的极化率均可从HMO法计算所得到的能级和波函数算出，而$\delta\alpha_r$、$\delta\beta_{rs}$则可以由积分参数的变化中算出。

实际上，PMO法已被人们用于诸如亲电芳香取代、亲核脂肪族取代、共轭烃的芳香性、周环反应、共轭体系中杂原子的取代效应、消除加成反应和π配合物稳定性的处理等方面。戴乾圜[8]运用PMO法以微扰分子轨道的离域能作参数，建立了PAH致癌的“双区理论”。

2.2.1.3 扩展的 Hückel 分子轨道法(EHMO 法或 EHT 法)

扩展的Hückel分子轨道法源于1952年Wolfsberg和Helmhoz的工作，后为Hoffmann所完善，在讨论烃类分子方面获得成功[9]。象HMO法一样，EHMO法明显地是经验性的。但EHMO法却是在Hückel近似的基础上扩充到计算全价电子的分子轨道方法，即EHMO法可用于计算非共轭分子。

1. 计算原理

我们知道，量子化学分子轨道法主要在于求解 Hartree-Fock-Roothaan 方程，其中最困难的是求出 Hartree-Fock 算符的矩阵元。由于 EHMO 法只考虑 Hartree-Fock 算符中的单电子贡献，因而其基本方程可表达为

$$\sum_{\nu}(H_{\mu\nu}-\varepsilon_i S_{\mu\nu})C_{\nu i}=0 \tag{2.65}$$

它在 HMO 法基础上，不忽略重叠积分S，而对H矩阵的矩阵元作如下处理：

$$H_{\mu\mu}=-I_{\mu} \tag{2.66}$$

式中 I_{μ} 是价轨道的电离势，可由实验获得。$H_{\mu\mu}$ 为H矩阵对角元。

$$H_{\mu\nu}=\frac{1}{2}K(H_{\mu\mu}+H_{\nu\nu})S_{\mu\nu}$$

其中，$H_{\mu\nu}$ 为矩阵非对角元，$S_{\mu\nu}$ 为重叠积分，K为比例常数，它的选取是 $1.5<K<2.0$，通常选1.75。非对角元$H_{\mu\nu}$ 的计算也有人建议选

$$H_{\mu\nu}=KS_{\mu\nu}\sqrt{H_{\mu\mu}\cdot H_{\nu\nu}} \tag{2.67}$$

或

$$H_{\mu\nu}=\frac{1}{2}(H_{\mu\mu}+H_{\nu\nu})S_{\mu\nu}(2-|S_{\mu\nu}|) \tag{2.68}$$

由于 EHMO 法没有考虑电子间相互作用，所以总电子能量等于电子轨道能量之和

$$E_{总}=\sum_i n_i\varepsilon_i \tag{2.69}$$

式中，n_i 为 i 轨道电子数，ε_i 为能级。

第 i 个分子轨道电荷集居数P_i 可由下式算出

$$\begin{aligned}P_i&=\sum_i n_i\Psi_i^2=\sum_i n_i\sum_{jj'}C_{ij}C_{ij'}(\varphi_j\varphi_{j'})\\&=\sum_i n_i\sum_{jj'}C_{ij}C_{ij'}S_{jj'}\\&=\sum_j\Big(\sum_i n_i\sum_{j'}C_{ij}C_{ij'}S_{jj'}\Big)\end{aligned} \tag{2.70}$$

EHMO 法在过去的一段时间里，是比较常用的量子化学方法，已有许多计算机程序发表。由于更精确的计算方法的建立，对此方法的应用已日趋减少。一般计算步骤往往是先计算重叠积分.建立H矩阵，再用求解广义特征值的计算方法得到能级和波函数，然后计算电荷集居数等指数。

在EHMO法计算中，经验参数的选择显然是重要的，其经验参数主要是价轨道的电离势。表 2-3 列出了EHMO 法的一些参数。

表 2-3　EHMO 参数表

原　　子	H	C	N	O	F
ζ*	1.00	1.625	1.95	2.275	2.60
Z**	1.00	3.25	3.90	4.55	5.20
$I(ns)$	13.60	21.43	27.50	35.30	38.24
$I(np)$		11.42	14.49	17.76	20.86

*ζ: STO(Slater 型轨道)轨道指数。

**Z: Slater 轨道中的 Z 参量（Z:有效核电荷数）。

其它含 d 轨道及 f 轨道的原子，在作EHMO 法计算时经验参数可从文献中查得，一般均已包含在EHMO 程序之中。

2．几何构型的表达

在EHMO法中，或在以后要叙述的自洽场法中，都必须给出一定的几何构型才能进行计算。

描述几何构型最简单的办法是把分子置于空间的直角坐标系中，把各原子核的坐标描述出来，一个有 n 个原子的分子应有 $3n$ 个这样的坐标。不过这 $3n$ 个坐标中有 3 个是用来描述分子的位置的，另外 3 个(对于直线型分子则为 2 个）是用来描述分子的取向的，剩下 的 $3n-6$（或 $3n-5$)个坐标才是用来描述分子的几何构型的。

实际上，几何构型用键长、键角和二面角来描述，会更符合化学的直观。这些键长、键角都可以看成一种广义坐标，通常称为内坐标。在通常使用的分子轨道计算程序中都有从内坐标转换成为直角坐标的计算功能。

内坐标对于每个原子以它的“键长”、“键角”和“二面角”来描述。这里首先要介绍二面角的概念。二面角是两个平面间的夹角。取第三个平面和所述的两个平面都正交，则交线的夹角就是所述的两个平面间的夹角。像直线间的夹角那样，二面角同样有 0 、锐角、直角、钝角、平角、反角等概念。也可以有负角。

分子内坐标的一种系统表达方法如下：对于每个原子，规定它的“键长”、“键角”和“二面角”作为坐标。这些坐标要关联到前面已确定了坐标的 3 个不在一直线上的原子。用L代表现在要规定坐标的原子的编号，I 、J 、K 代表所关联的原子的编号，称为关联序号。由于各原子的坐标是按其编号的次序逐一确定的，I 、J 和 K 都必须小于 L 。于是，原子 L 的“键长”定义为L与 I 的距离 R_{LI}，L 的“键角”定义为$\angle LIJ$，L 的“二面角”定义为平面 LIJ 与平面 IJK 间的夹角，

习惯上记为 LIJ/IJK。当角的绝对值小于 180°时,正负符号规定如下:若 L 处于平面 IJK 的右手螺旋正向(由 I 向 J,J 向 K,K 向 I 旋转),则 LIJ/IJK 取正值,在负向则取负值。图 2-3 中的 LIJ/IJK 是正的,LJI/JIK 则是负的。

从定义来看,所谓"键长"、"键角",无非是指几何关系,不一定要在这些地方存在着化学键。甚至关系序号所代表的也不一定是原子,而可以是所指处的一个位置。这种位置可称为虚设原子。

除了虚设 0 号的情况外,第一号原子 3 个关系序号皆不存在,也就不存在它的坐标;第二号原子只有键长;第三号原子只有键长和键角而无二面角;第四号原子三者都有。但在虚设了 0 号原子时,第一号原子有键长,第二号原子有键长和键角,第三号三者都有。

下面我们仍以胞嘧啶为例说明分子内坐标的一种系统表示方法。其中,键长和键角数据取自文献[9]。详见图 2-4 和表 2-4 (键长:nm;键角:度)。

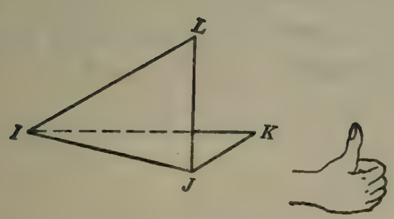


图 2-3

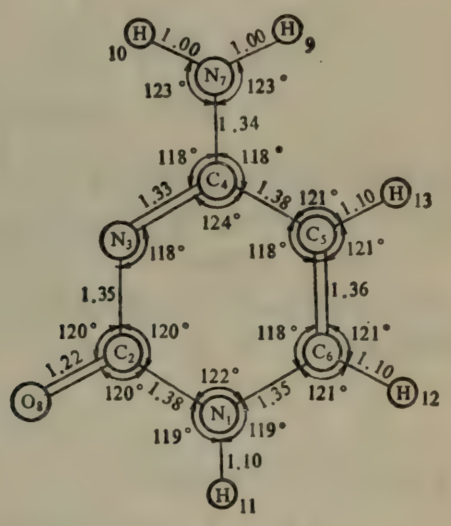


图 2-4

3. IEHT(电荷自洽 EHMO 法)

EHMO 法计算中,H 矩阵的近似主要取决于参数 I_μ,其误差也往往由于 I_μ 选择不同而显示出较大的差异。为了克服因 I_μ 选择的不同而造成的较大差异,科学工作者又作了如下修正:

$$H_{\mu\mu} = -I_\mu$$

$$I_\mu = I_\mu^0 + \alpha_\mu(Q_\mu^0 - Q_\mu) \tag{2.71}$$

其中,I_μ^0 系经验参数,Q_μ^0 为轨道 μ 上原来的电子数,而 Q_μ 则是从波函数系数 $C_{i\mu}$ 算出来的,α_μ 是适当选取的参数(>0)。

IEHT 方法现已被广泛应用于有机化合物和生物分子的计算(如 Pullman 等对核酸碱基电子分布和偶极矩的计算,参见 4.2.1.1)。对于含有过渡金属元素化合物的计算,则尤为需要电荷自洽。

表 2-4 胞嘧啶键长和键角数据

序号	键长	键角	二面角	I	J	K
1	0	0	0	0	0	0
2	0.138	0	0	0	0	0
3	0.135	120	0	2	1	0
4	0.133	118	0	3	2	1
5	0.138	124	0	4	3	2
6	0.136	118	0	5	4	3
7	0.134	118	180	4	5	6
8	0.122	120	180	2	3	4
9	0.100	123	0	7	4	5
10	0.100	123	0	7	4	3
11	0.110	119	180	1	2	3
12	0.110	121	180	6	1	2
13	0.110	121	180	5	6	1

2.2.2 半经验自洽场分子轨道法

这类方法的特点是忽略一些微分重叠，并用实验数据将积分参量化，以简化计算，节省时间并扩大计算范围。按照忽略微分重叠的程度又分为：CNDO法（全略微分重叠）、INDO 法（间略微分重叠）、NDDO 法（只忽略双原子微分重叠）以及 PRDDO 法（部分保留双原子微分重叠）。

在自洽场分子轨道理论中，对于闭壳层体系，经 LCAO 整理后，最终归结为求解 Hartree-Fock-Roothaan方程

$$\sum_{\nu}(F_{\mu\nu}-\varepsilon_i S_{\mu\nu})C_{\nu i}=0 \qquad (2.72)$$

式中，$F_{\mu\nu}$ 为 Hartree-Fock 算符矩阵元：

$$F_{\mu\nu}=H_{\mu\nu}+\sum_{\lambda\sigma}P_{\lambda\sigma}\left(\langle\mu\nu|\lambda\sigma\rangle-\frac{1}{2}\langle\mu\sigma|\lambda\nu\rangle\right) \qquad (2.73)$$

写成矩阵形式为

$$\boldsymbol{F}=\boldsymbol{H}+\boldsymbol{G} \qquad (2.74)$$

$\boldsymbol{H}$ 为 Hartree-Fock 矩阵的单电子部分，$\boldsymbol{G}$ 为双电子部分。

不难看出，LCAO 自洽场分子轨道计算的最困难和最费机时部分，是大量电子排斥积分的计算和处理，而且还要进行四中心，三中心积分的繁杂运算。为简化起见，在发展近似自洽场分子轨道的方案中，目前采用零微分重叠（ZDO）近似，已使得这类积分数目大大减少。

实践结果表明，$\Phi_\mu(1)\Phi_\mu(1)$ 型函数乘积的双电子积分比包含电子分布 $\Phi_\mu(1)\Phi_\nu(1)$（当 $\mu\neq\nu$）的积分值大得多，因而采用ZDO近似方法只考虑第一种积分

$$\Phi_\mu(1)\Phi_\nu(1)\mathrm{d}\tau=0\quad(\mu\neq\nu) \qquad (2.75)$$

于是，包含重叠分布的电子排斥积分小到可以忽略。实际上，四中心积分数组 $\langle\mu\nu|\lambda\sigma\rangle$ 已被简化成较容易计算的双中心积分数组 $\langle\mu\mu|\lambda\lambda\rangle$

$$\langle\mu\nu|\lambda\sigma\rangle=\langle\mu\mu|\lambda\lambda\rangle\delta_{\mu\nu}\delta_{\lambda\sigma} \tag{2.76}$$

式中，$\delta_{\mu\nu}\delta_{\lambda\sigma}$ 是 Kronecker δ 符号。在分子轨道归一化中，忽略相应的重叠积分

$$s_{\mu\nu}=\int\Phi_\mu(1)\Phi_\nu(1)\mathrm{d}\tau=\delta_{\mu\nu} \tag{2.77}$$

经 ZDO 近似后，闭壳层分子 LCAO 系数的 HFR 方程可写成简化形式

$$\sum F_{\mu\nu}\cdot C_{\nu i}=\varepsilon_i\cdot C_{\nu i} \tag{2.78}$$

由此可见，经 ZDO 近似后，即可将求解广义本征值问题简化为一般本征值问题。

目前常用的半经验自洽场分子轨道法都是建立在这个基础之上的，不同等级的近似。其主要区别在于处理电子排斥积分中应用 ZDO 近似的程度。以下主要讨论量子生物学中 较常用的 CNDO/2、INDO 和 MNDO 法。

2.2.2.1 全略微分重叠法(CNDO 法)

1．计算原理

CNDO 法是 ZDO 近似用于全部原子轨道的近似方法，亦即完全采用(2.75)、(2.76)式的近似。这样，Fock 矩阵元可写成

$$F_{\mu\mu}=H_{\mu\mu}-\frac{1}{2}P_{\mu\mu}\langle\mu\mu|\mu\mu\rangle+\sum_\lambda P_{\lambda\lambda}\langle\mu\mu|\lambda\lambda\rangle \tag{2.79}$$

$$F_{\mu\nu}=H_{\mu\nu}-\frac{1}{2}P_{\mu\nu}\langle\mu\mu|\nu\nu\rangle \tag{2.80}$$

由此二式可知，在 CNDO 近似中，只剩下库仑型积分的集合。如果设轨道 λ 属于原子 A，轨道 μ 属于原子 B，则可简化为每一对原子一个数值

$$\langle\lambda\lambda|\mu\mu\rangle=\gamma_{AB} \tag{2.81}$$

假设$\langle\lambda\lambda|\mu\mu\rangle$只与属什么原子有关，而不论它属于什么轨道。于是，我们便可取最简单的 s 轨道来计算

$$\begin{aligned}\langle\lambda\lambda|\mu\mu\rangle=\gamma_{AB}&=\iint\varphi_\lambda(1)\varphi_\lambda(1)\frac{1}{r_{12}}\varphi_\mu(1)\varphi_\mu(1)\mathrm{d}\tau_1\mathrm{d}\tau_2\\&=\iint s_A^2(1)\,\frac{1}{r_{12}}s_B^2(1)\mathrm{d}\tau_1\mathrm{d}\tau_2\end{aligned} \tag{2.82}$$

这样，(2.79)、(2.80)可写作

$$F_{\mu\mu}=H_{\mu\mu}-\frac{1}{2}P_{\mu\mu}\gamma_{AA}+\sum_B P_{BB}\gamma_{AB}\quad \Phi_\mu\text{ 属于 A} \tag{2.83}$$

$$F_{\mu\nu}=H_{\mu\nu}-\frac{1}{2}P_{\mu\nu}\gamma_{AB}\quad \Phi_\mu\text{ 属于 A},\Phi_\nu\text{ 属于 B} \tag{2.84}$$

式(2.83)中，P_{BB} 表示受 B 原子束缚的总电荷密度

$$P_{BB}=\sum_\lambda^B P_{\lambda\lambda} \tag{2.85}$$

式中求和遍及 B 上全部原子轨道。

下一步是应用一系列有关近似于分子实的 Hamilton 矩阵元$H_{\mu\nu}$。对角元$H_{\mu\mu}$可方便地分离

为单中心和双中心贡献，若 Φ_μ 属于A，则可写作

$$H_{\mu\mu}=U_{\mu\mu}-\sum_{B(\neq A)}\langle\mu|\boldsymbol{V}_B|\mu\rangle \tag{2.86}$$

同理，考虑原子 A 的不同原子轨道 Φ_μ 和 Φ_ν 间非对角实矩阵元 $H_{\mu\nu}$ 也可分成两部分

$$H_{\mu\nu}=U_{\mu\nu}-\sum_{B(\neq A)}\langle\mu|\boldsymbol{V}_B|\nu\rangle \quad \Phi_\mu、\Phi_\nu \text{ 属于 A} \tag{2.87}$$

式中，$U_{\mu\mu}$ 可以根据原子数据半经验地获得，后一项计算所采取的近似方法，应与处理双电子积分的近似一致，于是忽略 A 原子上的单电子微分重叠，亦即 $\langle\mu|\boldsymbol{V}_B|\nu\rangle$ 取为零，则有：

$$\langle\mu|\boldsymbol{V}_B|\mu\rangle=V_{AB} \tag{2.88}$$

作为这些近似的结果，可得

$$H_{\mu\mu}=U_{\mu\mu}-\sum_{B(\neq A)}V_{AB} \quad \Phi_\mu \text{ 属于 A} \tag{2.89}$$

$$H_{\mu\nu}=0 \quad \Phi_\mu \neq \Phi_\nu，\text{两者均属 A} \tag{2.90}$$

在 CNDO 近似中，取双原子非对角实矩阵元正比于相应的重叠积分

$$H_{\mu\nu}=\beta^0_{AB}s_{\mu\nu} \quad \Phi_\mu \text{ 属于 A，} \Phi_\nu \text{ 属于 B} \tag{2.91}$$

式中，β^0_{AB} 只与 A、B 原子的本质有关，$s_{\mu\nu}$ 重叠积分的计算跟 EHMO 法中重叠积分的计算相同。

应用这些近似后，Fock 矩阵元简化为下列简单形式(Φ_μ 属于 A，Φ_ν 属于B)

$$F_{\mu\mu}=U_{\mu\mu}+(P_{AA}-\frac{1}{2}P_{\mu\mu})\gamma_{AA}+\sum_{B(\neq A)}(P_{BB}\gamma_{AB}-V_{AB}) \tag{2.92}$$

$$F_{\mu\nu}=\beta^0_{AB}s_{\mu\nu}-\frac{1}{2}P_{\mu\nu}\gamma_{AB} \quad \mu\neq\nu \tag{2.93}$$

式中

$$\beta^0_{AB}=\frac{1}{2}K(\beta^0_A+\beta^0_B) \tag{2.94}$$

β^0_A 和 β^0_B 分别为 A 原子与 B 原子的成键参量，K 是可调整的常数。一般，$K=1$，A 和(或) B 属于第三周期元素的原子时，$K=0.75$。

由上述可知，一个 CNDO 计算的完成，需要重叠积分 $s_{\mu\nu}$，$U_{\mu\mu}$、V_{AB}，电子排斥积分 γ_{AB}。迄今为止，已经建立了多种方案来获得这些参量。

在量子生物学研究中，CNDO/2 法* 已被广泛采用。它是经 Pople、Segal 等人实现的[10]。在 CNDO/2 方案中，不再个别地计算电子——实势能积分 V_{AB}，而是将 B 原子核实视为点电荷，借助下式把 V_{AB} 与电子排斥积分联系起来。

* 一个 CNDO 计算的完全确定，需要重叠积分 $s_{\mu\nu}$，实 Hamilton 量元素 $U_{\mu\mu}$，V_{AB}，电子排斥积分 γ_{AB} 和成键参量 β^0_{AB} 的数值。已经建议了两种方法来得到它们，称为 CNDO/1 和 CNDO/2。第二种方法更成功一些，且已被广泛应用。

$$V_{AB}=Z_B r_{AB} \tag{2.95}$$

采用上述经验公式计算双中心积分将大大简化计算，且经研究证实，在一些情况下给出的结果与实验符合较好。

CNDO/2 方案中另一个参量化是 $U_{\mu\mu}$。在 CNDO/1 方案中，$U_{\mu\mu}$ 依据对应的平均原子态的电离势 I_μ，由关系式

$$-I_\mu=U_{\mu\mu}+(Z_A-1)r_{AA} \tag{2.96}$$

求得，其中原子轨道 Φ_μ 属于 A 原子。另一方法应用原子的电子亲合势 A_μ，相应的关系式为

$$-A_\mu=U_{\mu\mu}+Z_A r_{AA} \tag{2.97}$$

在一个概括性更大的近似分子轨道理论中，希望同时说明一个原子轨道得失电子的两种倾向。CNDO/2 法系将(2.96)和(2.97)式平均而得到参数

$$-\frac{1}{2}(I_\mu+A_\mu)=U_{\mu\mu}+\left(Z_A-\frac{1}{2}\right)r_{AA} \tag{2.98}$$

这里，$(I_\mu+A_\mu)/2$ 是 μ 原子轨道的电负性，可由实验获得。

按照(2.92)、(2.93)、(2.95)和(2.98)式，CNDO/2 方法中 Fock 矩阵元可写成

$$F_{\mu\mu}=-\frac{1}{2}(A_\mu+I_\mu)+\left[(P_{AA}-Z_A)-\frac{1}{2}(P_{\mu\mu}-1)\right]r_{AA}$$

$$+\sum_{B(\neq A)}(P_{BB}-Z_B)r_{AB} \tag{2.99}$$

$$F_{\mu\nu}=\frac{1}{2}K(\beta_A^0+\beta_B^0)s_{\mu\nu}-\frac{1}{2}P_{\mu\nu}r_{AB} \tag{2.100}$$

其密度矩阵和总能量可写成

$$P_{\mu\nu}=2\sum_i C_{i\mu}C_{i\nu} \tag{2.101}$$

$$\varepsilon_{总}=\varepsilon_{单原子}+\varepsilon_{双原子}$$

$$=\sum_A \varepsilon_A+\sum_{A<B}\varepsilon_{AB} \tag{2.102}$$

$$\varepsilon_A=\sum_\mu^A P_{\mu\mu}U_{\mu\mu}+\frac{1}{2}\sum_\mu^A\sum_\mu^B\left(P_{\mu\mu}P_{\nu\nu}-\frac{1}{2}P_{\mu\nu}^2\right)r_{AA} \tag{2.103}$$

$$\varepsilon_{AB}=\sum_\mu^A\sum_\mu^B\left(2P_{\mu\nu}\beta_{\mu\nu}-\frac{1}{2}P_{\mu\nu}^2 r_{AB}\right)+Q_AQ_BR_{AB}^{-1} \tag{2.104}$$

其中，Q_A、Q_B 分别为 A、B 原子的价电子电荷，R_{AB} 为两原子间距。

至此，CNDO/2 法便可方便地通过计算机程序的编制来实现。(2.99)、(2.100)十分简洁地表明了 CNDO/2 法的自洽特征。

在实际计算时，首先应计算重叠积分 $s_{\mu\nu}$，双中心库仑积分 r_{AB}。对于初始的 Fock 矩阵，按常规取 Hückel 近似，即轨道 Φ_μ 包含一个电子（$P_{\mu\mu}=1$），且全部原子净电荷为零（$P_{AA}=Z_A$，

$P_{BB}=Z_B$)，则对角元 $F_{\mu\mu}$ 简化为 $-\frac{1}{2}(I_\mu+A_\mu)$。$F_{\mu\nu}$ 简化为 $\frac{1}{2}s_{\mu\nu}(\beta_A^0+\beta_B^0)K$，即可求解 HFR 方程，依据第一次求解得到的波函波计算密度矩阵，并用(2.99)、(2.100)式计算新的 Fock 矩阵元，如此循环进行，直至叠代收敛为止。

2. CNDO/2 参数

在表 2-5 和表 2-6 中，我们给出了第一至第三周期元素的 CNDO/2 参数，其它参数可从有关文献中查到，这里不再一一列出。

表 2-5　第一、二周期元素 CNDO/2 参数(eV)

原　　子	H	Li	Be	B	C	N	O	F
$\frac{1}{2}(I_s+A_s)$	7.176	3.106	5.946	9.594	14.051	19.316	25.390	32.272
$\frac{1}{2}(I_p+A_p)$		1.258	2.563	4.001	5.572	7.275	9.111	11.080
$-\beta_A^0$	9	9	13	17	21	25	31	39

表 2-6　第三周期元素 CNDO/2 参数(eV)

原　　子	Na	Mg	Al	Si	P	S	Cl
$\frac{1}{2}(I_{3s}+A_{3s})$	2.804	5.125	7.771	10.033	14.033	17.650	21.591
$\frac{1}{2}(I_{3p}+A_{3p})$	1.302	2.052	2.995	4.133	5.464	6.989	8.708
$\frac{1}{2}(I_{3d}+A_{3d})$	0.150	0.162	0.224	0.337	0.500	0.713	0.977
$-\beta_A^0$	7.720	9.447	11.301	13.065	15.070	18.150	22.330

选好了 CNDO/2 参数，就使这一方案得以具体实现。在这里，矩阵元的准备工作比从头计算简单得多，它只须计算重叠积分及少量排斥积分以构成 Fock 矩阵。这些积分的计算量只占全部计算量中的一小部分。

3. 生物分子的 CNDO/2 计算

CNDO/2 法已经广泛地应用于量子生物学研究。迄今为止，比较集中的是用于药物分子、氨基酸多肽、核酸碱基、核苷酸，甚至更大的分子的理论计算以及结构活性相关的分析。因而，我们在这一节里用了较大的篇幅予以介绍。按照作者等的经验，这样做将是有益的。纵观这一方法在上述领域应用的实践可以看出，CNDO/2 法在量子生物学中应用的主要特点是：计算分子中所含原子的种类比较单一(绝大部分限于第三周期以前的元素，对于过渡金属配位化合物的计算也有报导)[11]，分子规模大，计算与实验结果一般比较吻合。

从目前国内计算机应用状况考虑，生物分子的半经验法计算，着重是要解决叠代收敛和加速收敛的问题。前者是为完成计算的自洽克服叠代振荡，后者是考虑减少叠代次数节约机时。从解决振荡问题看，各种实用程序大都没有引进叠代因子强迫收敛的技术。根据已经计算过的大分子化合物的分析来看[12]，若分子中含 S—S、O—O、N—N 键等，以及大分子中某一空间聚

集较多的O、S、N和P等杂原子时,叠代就可能出现振荡现象。尤其应该注意的是,上述这些化合物即使能量自洽,且符合程序给定的收敛限(如$\Delta E<10^{-6}$Hartree),这些化合物的杂原子净电荷也有可能产生振荡现象。为克服振荡,这时可以用叠代因子再叠代 2—3 次,一般即可获得满意的结果。在具体进行大分子 CNDO/2 或其它自洽场计算时,可将计算先进行 5 次叠代,并将计算结果存入计算机,再分析打印的 5 次叠代进程,即可看出能量是否趋于收敛。若出现振荡,则可应用叠代因子,再继续往下叠代。此时叠代计算的 Fock矩阵元等均取自事先存入计算机的数值,即第五次叠代的结果。这样一来,既保证了叠代收敛,又可节省计算机时。

这里,尤其需要补充说明的是,宣建城和江寿平[13]提出的一个新的叠代公式——"对角位移公式",可以解决收敛和减少叠代次数。该公式为:

$$A^{(k)}=\alpha^{(k)}\cdot A^{*}+C^{(k-1)}\cdot X^{(k)}\cdot C^{(k-1)T}$$

式中 A^{*} 为第 $k-1$ 次叠代按简单叠代法得到的 Fock 矩阵。$C^{(k-1)}$为第 $k-1$ 次叠代得到的波函数,$C^{(k-1)T}$ 是它的转置矩阵,$\alpha^{(k)}$ 为系数,$X^{(k)}$ 为一对角方阵。利用这一公式已对 AMP、ADP 和 ATP 等分子成功地进行了 CNDO 计算,效果显著。

2.2.2.2 修改的忽略双原子微分重叠法(MNDO 法)

假如在化简 HFR 方程时,只对 Φ_μ 和 Φ_ν 属于不同原子的原子轨道才使用ZDO近似((2.75)式),而这一级近似的主要特点是保留了偶极-偶极相互作用。因为此时包含了$\langle S_AP_A|S_AP_A\rangle$型积分,大体上正比于 R^{-3}(R 表示核间距)。像这种只忽略双原子微分重叠的近似方法称为 NDDO 法。

NDDO 法虽然忽略了全部四中心和三中心积分$\langle\mu\nu|\lambda\sigma\rangle$,但是所保留的单中心和双中心积分数目仍然很多,因而这种方法就比较难以过渡到半经验的水平上去。因为对双中心积分大都需用 Slater 型轨道作直接计算,且程序繁杂,因而只能计算一些小分子,这就影响到这种方法在量子生物学研究中的实际应用。于是,针对 NDDO 法的一些不足之处,Dewar 及其合作者提出了改进的忽略双原子微分重叠法。该方法采用半经验模型处理双中心电子排斥积分,这种方法称作 MNDO 法。

在 NDDO 法中 Fock 矩阵元为

$$F_{\mu\mu}^{AA}=U_{\mu\mu}+\sum_{B}V_{\mu\mu}^{B}+\sum_{\nu}^{A}P_{\mu\nu}[\langle\mu\mu|\nu\nu\rangle-\frac{1}{2}\langle\mu\nu|\mu\nu\rangle]$$
$$+\sum_{B}\sum_{\lambda\sigma}^{B}P_{\lambda\sigma}\langle\mu\mu|\lambda\sigma\rangle \tag{2.105}$$

$$F_{\mu\nu}^{AA}=\sum_{B}V_{\mu\nu}^{B}+\frac{1}{2}P_{\mu\nu}[3\langle\mu\nu|\mu\nu\rangle-\langle\mu\mu|\nu\nu\rangle]$$
$$+\sum_{B}\sum_{\lambda\sigma}^{B}P_{\lambda\sigma}\langle\mu\nu|\lambda\sigma\rangle \tag{2.106}$$

$$F_{\mu\lambda}^{AB}=\beta_{\mu\lambda}-\frac{1}{2}\sum_{\mu}^{A}\sum_{\sigma}^{B}P_{\mu\sigma}\langle\mu\nu|\lambda\sigma\rangle \tag{2.107}$$

从这些公式可知，有待计算确定的量是：

$U_{\mu\mu}$，单中心单电子能量，表示 A 原子中 Φ_μ 轨道的动能及其在 A 原子核实场中的势能。

$\langle\mu\mu|\nu\nu\rangle=g_{\mu\nu}$，单中心双电子排斥积分，库仑积分。

$\langle\mu\nu|\mu\nu\rangle=h_{\mu\nu}$，交换积分。

$\beta_{\mu\lambda}$，双中心单电子核实共振积分。

$V^{B}_{\mu\nu}$，A 原子的电荷分布 $\Phi_\mu\Phi_\nu$ 与B 原子核实的吸引能。

$\langle\mu_A\nu_A|\lambda_B\sigma_B\rangle$，双中心双电子排斥积分。

现在，我们来讨论 Dewar 提出的 MNDO 法。该方法的主要特点是，把双原子的电荷分布 $\Phi^A_\mu\Phi^A_\nu$ 与 $\Phi^B_\lambda\Phi^B_\sigma$ 间的相互作用看成是两堆电荷分布的多极矩相互作用之和

$$\langle\mu_A\nu_A|\lambda_B\sigma_B\rangle=\sum_{l_1}\sum_{l_2}\sum_{m}[M^{A}_{l_1m_1},M^{B}_{l_2m_2}] \tag{2.108}$$

其 l,m 分别表示多极矩 M_{lm} 的级数和取向。这种近似展开必须保证多极矩-多极矩相互作用应有好的渐近行为。当 $R_{AB}\to\infty$ 时，它们应收敛到经典相互作用；当 $R_{AB}=0$ 时，它们则应当重现相应单中心积分的半经验数值。Dewar 等将多极矩-多极矩相互作用近似地选为

$$[M^{A}_{l_1m_1},M^{B}_{l_2m_2}]=\frac{e^2}{2^{l_1+l_2}}\sum_{i=1}^{2^{l_1}}\sum_{j=1}^{2^{l_2}}f_1(R_{ij}) \tag{2.109}$$

$f_1(R_{ij})$ 表示当 $R_{AB}\to\infty$ 和 $R_{AB}\to 0$ 时的半经验表达式。数值为$\frac{e}{2^l}$的点电荷相距 D_l 时的多极矩记作$[M_{lm}]$。对于仅含 sp 价壳层的第二周期元素的原子，可得

$$D_1=\frac{5}{3^{\frac{1}{2}}}\frac{(4\zeta_{2s}\zeta_{2p})^{5/2}}{(\zeta_{2s}+\zeta_{2p})^6} \tag{2.110}$$

$$D_2=\left(\frac{3}{2}\right)^{1/2}\zeta_{2p}^{-1} \tag{2.111}$$

式中，ζ_{2s} 和 ζ_{2p} 是 2 s 和 2 p 轨道的指数。对于半经验函数 $f_1(R_{ij})$选成

$$f_1(R_{ij})=[R^2_{ij}+(\rho^A_{l_1}+\rho^B_{l_2})^2]^{-\frac{1}{2}} \tag{2.112}$$

对于 sp 基的原子而言，只用三个相加项 ρ_l；同时 $l=0,1,2$ 分别表示单极、偶极、四极的性质。两个单极相互作用给出 g_{ss}，两个偶极相互作用极限为 h_{sp}，两个四极矩则给出极阻 h_{pp}。因而 $\rho_0=e^2/2\,g_{ss}$、ρ_1 和 ρ_2 分别由数值计算给出。这样，就可根据公式(2.109)、(2.110)、(2.112)，用确定的电荷距离 D_l 和相加项 ρ_l 即可表示出 MNDO 法中的电子排斥积分。

综上所述，MNDO 法只包含原子型参数 $U_{\mu\mu}$，ζ，β^A_{μ} 和 α_A，它们经过细致的参数化后的最优值列于表 2-7 中。于是，利用这些参数和上述原理即可形成 Fock 矩阵，再利用计算机进行通常的自洽叠代，最后计算理论指数。

由于 MNDO 法提供了对完整的 Hartree-Fock 矩阵的最相近的处理，考虑了高级矩的不同取向。对涉及化学键方向性的问题，方法中在 $\beta_{\mu\nu}$、双中心排斥积分、核实——电子吸引积分的近似公式中都作了考虑，故它在处理分子中成键方向性，计算不饱和分子生成热，几何优化所得的构象，计算出的分子轨道次序与光电子能谱数据的吻合程度等方面，均优于修改的 INDO 法，即 MINDO/3 法。所以尽管 MNDO 法的计算时间大约比 MINDO/3 法多 20%，但仍旧被看做是

表 2-7 最优MNDO参数

参数 \ 原子	H	B	C	N	O	F
U_{ss} (eV)	−11.906276	−34.547130	−52.279745	−71.932122	−99.64309	−131.071548
U_{pp} (eV)		−23.121690	−39.205558	−57.172319	−77.797472	−105.782137
ζ (原子单位)	1.331967	1.506801	1.787537	2.255614	2.699905	2.848487
β_s(eV)	−6.989064	−8.252054	−18.985044	−20.495758	−32.688082	−48.290460
β_p(eV)		−8.252054	−7.934122	−20.495758	−32.688082	−36.508540
α(埃$^{-1}$)	2.544134	2.134993	2.546380	2.861342	3.160604	3.419661

目前最精确的半经验分子轨道法而被越来越广泛的使用。在量子生物学研究中，MNDO 法日益受到重视，它被科学工作者用于化学致癌机理、生物分子的电子结构和量子药理学研究，取得了一些有意义的成果[14-16]。特别是 QCPE(量子化学程序交换) 提供的几何优化的 MNDO 程序，可以首先通过选定的键长、键角、二面角进行几何优化计算，获得稳定构象后再进行最后迭代计算。这一方法可以看作是当前量子生物学中除了 ab initio(从头计算法)之外，最有前途的理论方法。

2.2.3 从头计算法(ab initio 法)

从头计算法(ab initio calculation 法) 隶属于分子轨道理论的计算方法。它以三个近似，即非相对论近似，Born-Oppenheimer 近似和轨道近似(参见 2.1.2)为出发点，不借助于任何经验参数而严格求解HFR方程。

从头计算法是非经验求解多电子体系问题的量子理论全电子计算方法。由于理论上的严格性和计算结果的准确可靠性，使它在各种量子化学计算方法中居于主导地位。由于从头计算法所获得的各类体系(如原子、离子、分子、原子簇及化学反应体系等)的电子微观运动状态及其各种量子化学指数，能合理地解释与预测原子间的键合、分子的结构、化学反应的过程、物质的性质、生物分子的活性及有关实验事实，因此，它已成为理论化学研究的一种必不可少的理论工具，并且已经广泛地应用于化学的各个分支学科，并渗透到生物学和医学等领域。

1. 计算原理[17]

概括地讲，从头计算方法实质上是以分子轨道理论为基础的，它以 Hartree-Fock-Roothaan 方程为出发点，在适当地选取基函数后，计算各种所需要的积分，然后进入自洽求解。

按照三个基本近似，可以将单电子波函数 Ψ_i 写成某种固定基函数的线性组合。在采用分子轨道——原子轨道线性组合基函数的情况下，对体系的能量表达式进行变分处理，就得到著名的 HFR 方程(详细推导可参考"量子化学"专著)。对闭壳层体系而言，HFR 方程具有形式：

$$FT = STE \tag{2.113}$$

式中 T 为原子轨道组合系数矩阵，S 为重叠积分矩阵，E 为能级，F 为 Fock 矩阵，其矩阵元为

$$F_{ij} = h_{ij} + G_{ij}$$

$$=\langle\Phi_i|\hat{h}|\Phi_j\rangle+\sum_{k,l}P_{k,l}[2\langle\Phi_i\Phi_k|\hat{g}|\Phi_j\Phi_l\rangle-\langle\Phi_i\Phi_k|\hat{g}|\Phi_l\Phi_j\rangle] \tag{2.114}$$

取 T 矩阵中的某一列矩阵 C_k 表示分子轨道，则可写成广义本征方程

$$FC_k=\varepsilon_k sC_k \quad (k=1,2,\cdots,n) \tag{2.115}$$

用求解广义本征值的方法即可获得本征值 ε_k，而本征向量即为分子轨道系数。

由上式解得的每一个分子轨道都容纳两个自旋相反的电子，所以它被称为自旋限制性方程。

从 Fock 矩阵元中不难看出，由于它包含 $P_{k,l}$ 这一待求系数，因此 Roothaan 方程须用自洽场方法求解。我们先假设一组试探分子轨道 $C_1^{(0)}$ ，$C_2^{(0)}$ ，…，$C_m^{(0)}$，定义出矩阵 $T^{(0)}$，以形成密度矩阵 $D^{(0)}$，而后按式(2.122)建立 Fock 矩阵 $F^{(0)}$。求解 $F^{(0)}C^{(1)}=\varepsilon^{(1)}SC^{(1)}$，就可得到广义本征能量 $\varepsilon_1^{(1)},\varepsilon_2^{(1)},\cdots,\varepsilon_m^{(1)}$ 及相应广义本征向量 $C_1^{(1)},C_2^{(1)},\cdots,C_m^{(1)}$。然后，由这组本征向量组成第二次试探分子轨道，再开始第二次叠代，周而复始直至相邻两次叠代所求得的轨道波函数相一致(误差应在给定范围内)。

以上我们就闭壳层体系作了扼要的介绍。对于开壳层体系，其自旋限制性方程就比较复杂了。一般常用自旋非限制性的 HFR 方程来求解开壳层体系，它可表达成联立的矩阵方程组：

$$F^\alpha T^\alpha=ST^\alpha E^\alpha \tag{2.116}$$

$$F^\beta T^\beta=ST^\beta E^\beta \tag{2.117}$$

式中有两个 Fock 矩阵，其元素为

$$F_{ij}^\alpha=h_{ij}+\sum_{k,l}[(P_{k,l}^\alpha+P_{k,l}^\beta)\langle\Phi_i\Phi_k|\hat{g}|\Phi_j\Phi_l\rangle-P_{k,l}^\alpha\langle\Phi_i\Phi_k|\hat{g}|\Phi_l\Phi_j\rangle] \tag{2.118}$$

$$F_{ij}^\beta=h_{ij}+\sum_{k,l}[(P_{k,l}^\alpha+P_{k,l}^\beta)\langle\Phi_i\Phi_k|\hat{g}|\Phi_j\Phi_l\rangle-P_{k,l}^\beta\langle\Phi_i\Phi_k|\hat{g}|\Phi_l\Phi_j\rangle] \tag{2.119}$$

求解的方法亦须借助类似的自洽场叠代方法，给出 P^α 和 P^β 的一个假设的初值，就可形成两个 Fock 矩阵 F^α 和 F^β，解出联立矩阵方程组，得两组本征向量，分别取出本征能量最低的 $m^\alpha=p$ 和 $m^\beta=q$ 个被占据轨道，形成一个新的密度矩阵，循环重复以上处理过程，直至 α 和 β 两组分子轨道前后两次叠代的误差分别在允许的精度范围，即自洽为止。

无论是限制性方程或是非限制性方程，在求得自洽后的分子轨道系数后，便可依次而计算有关分子电子结构的一些指数，如偶极矩、Mulliken 集居数分析和分子静电势(参见后续章节)等。

(1) Gauss 函数(或 Gauss 型轨道，GTO)

我们知道，虽然 Slater 型原子轨道(STO)能较好地反映原子或分子中电子所处的势场，但若把它应用于从头计算，则将会给分子积分的计算带来很大麻烦。因此，在从头计算法中常以 Gauss 函数作为基函数，以便使各种多中心积分容易计算，从而得到解析表达式。

Gauss 函数共有五个参数。它通常采用由Boys提出的用归一化Gauss函数作为径向函数，由

此得到的基函数称为 Gauss 型轨道基函数，简称 GTO。在量子化学计算中，GTO 一般都表示成直角坐标系中的实函数形式，以坐标原点为中心的 GTO 可写成

$$\chi_A = N_A x_A^l y_A^m z_A^n e^{-\alpha r_A^2} \tag{2.120}$$

式中 N_A 是归一化系数，A 指明坐标系原点，x_A、y_A、z_A 是某一空间点的笛卡尔坐标，r_A 是原点到空间点的距离，l,m,n 分别为相应原子轨道的轨道角量子数。

Gauss 函数集合是完全集合，它包含了 Hermite 函数集合。Gauss 函数彼此之间是非正交的，它们有下列性质：

① 两个不同中心的 Gauss 函数相乘可以表达为另一个单中心的 Gauss 函数的线性组合

$$\chi(A,\alpha_1,l_1,m_1,n_1)\chi(B,\alpha_2,l_2,m_2,n_2) = \sum_{l,m,n} C_{lmn}\chi(P,\beta,l,m,n) \tag{2.121}$$

新的 Gauss 函数的中心 P 点定义为中心 A 和中心 B 的质心坐标

$$P = \frac{\alpha_1 A + \alpha_2 B}{\alpha_1 + \alpha_2} \tag{2.122}$$

展开系数 C_{lmn} 可以根据 A、B、P 三个中心到同一空间点的矢量关系求得。

② Gauss 函数对笛卡尔坐标 x、y、z 的积分可分离变量。

③ Gauss 函数积分变换公式：对变量 x 的幂函数积分，可通过 Gauss 函数变换为新参量的 Gauss 函数

$$\int_{-\infty}^{\infty} e^{iyx} x^n e^{-ax^2}\,\mathrm{d}x = i^n\left(\frac{\pi}{a}\right)^{\frac{1}{2}}\left(\frac{1}{2\sqrt{a}}\right)^n H_n\left(\frac{y}{2\sqrt{a}}\right)e^{-(y^2/4a)} \tag{2.123}$$

其中 $H_n(z)$是 Hermite 函数。

(2) 多中心积分的计算

欲计算的多中心积分有：

① 重叠积分

$$s_{AB} = \int \chi(A,\alpha_1,l_1,m_1,n_1)\chi(B,\alpha_2,l_2,m_2,n_2)\mathrm{d}\tau \tag{2.124}$$

② 动能积分

$$\begin{aligned} KE &= \int \chi(A,\alpha_1,l_1,m_1,n_1)\left(-\frac{1}{2}\nabla^2\right)\chi(\beta,\alpha_2,l_2,m_2,n_2)\mathrm{d}\tau \\ &= -\frac{1}{2}\int \chi(A,\alpha_1,l_1,m_1,n_1)\left(\frac{\partial^2}{\partial x_B^2}+\frac{\partial^2}{\partial y_B^2}+\frac{\partial^2}{\partial z_B^2}\right) \\ &\quad \cdot \chi(B,\alpha_2,l_2,m_2,n_2) \end{aligned} \tag{2.125}$$

③ 核吸引能积分

以 C 点为中心的核对电子云的吸引而产生的积分

$$NAI = \int \chi(A,\alpha_1,l_1,m_1,n_1)\frac{1}{r_C}\chi(B,\alpha_2,l_2,m_2,n_2)\mathrm{d}\tau \tag{2.126}$$

④ 电子排斥能积分

$$ERI=\int\chi(A,\alpha_1,l_1,m_1,n_1)\chi(B,\alpha_2,l_2,m_2,n_2)\frac{1}{r_{12}}$$
$$\cdot\chi(C,\alpha_3,l_3,m_3,n_3)\chi(D,\alpha_4,l_4,m_4,n_4)\mathrm{d}\tau_1\mathrm{d}\tau_2 \tag{2.127}$$

在以上各类能量积分中，动能积分是双中心积分，核吸引能积分是三中心积分，而电子排斥能积分是四中心积分。

现在，让我们回忆前面所讨论过的各种半经验方法，不难归纳出：在 HMO 和 EHMO 方法中，由于引入了有效单电子 Hamilton 算符，所有能量积分都是单电子积分，并且对这类单电子能量积分不做具体计算，而是近似为经验参数。在 CNDO 方法中，虽然对单电子能量积分遇到三中心，对双电子能量积分遇到四中心问题，但在采取了 ZDO 近似后，单电子能量积分代之以经验参数，双电子能量积分转化为单中心、双中心积分进行计算。由此看来，半经验方法迴避了计算多中心积分的问题。

与半经验法不同，从头计算法不引入任何经验参数，各类积分都作严格计算。但是，它利用 GTO 的特性，将分子积分中的多中心积分化为单中心积分的解析表达式，使计算简化。文献[18]给出了采用 GTO 计算(2.114)、(2.118)和(2.119)式中各项分子积分的解析表达式。

在目前通用的从头计算法电子计算机程序中，往往采用 STO-GTO 系基函数。这是由于采用 GTO 虽然具有计算迅速、准确度高的优点，但是直接作为基函数线性组合成分子轨道的GTO本身与电子运动的真实图象相差甚远。在按照分子轨道讨论分子中原子间成键的性质时，又不便将 GTO 与通常化学概念中的原子轨道相联系。为了弥补这一不足，因此常常用 GTO 的线性组合近似 STO 的办法。这种基组，以 STO 为 SCF 的基函数，而每个 STO 又用若干个 GTO 来逼近(一般是 3—6 个)。解出的分子轨道仍用 STO 的线性组合表示。这样，GTO 仅作为中间数学处理方法，而不再作为原子轨道，我们将其称为 STO-GTO 系，其一般表达式为

$$\Phi_{nlm}(\xi,r)=\sum_{K=1} d_{n_k lm}\chi_{n_k lm}(\alpha n_k;r^2) \tag{2.128}$$

式中 $\Phi_{n,l,m}(\xi,r)$表示 STO，$\chi_{n_k lm}(\alpha n_k;r^2)$为 GTO，$d_{n_k lm}$为线性组合系数。

经展开后可得

$$\Phi_k(\xi,r)=f_k(x,y,z)\sum_i d_{ki}N_i(k,\xi)\mathrm{e}^{-\xi^2\alpha_i r^2} \tag{2.129}$$

式中 d_{ki} 为展开系数，$N_i(k,\xi)$为 $\chi_i(\xi^2\alpha_i;r^2)$的归一化系数，它与 α_i，f_k，ξ 有关

Huzinaga[19]、Pople 等[20]系统地研究了 STO 向 GTO 展开的问题，先后发表了至第五周期原子轨道的若干 STO-GTO 展开基组的组合系数 d_i 和轨道指数 α_i，已为各国量子化学工作者所采用。我国陈凯先等[21]采用无约束的多因素优化方法——Sagent 改进了的 Powell 共轭方向法和抛物线拟合的一维寻找法，确定 GTO 组合系数 d_i 和轨道指数 α_i，第一次发表了所获得的第六周期的 STO-GTO 展开的全部数据。

2．大分子的从头计算

由于从头计算法计算工作量很大，特别是其中最难计算的双电子积分数目和所用的轨道数目(此数目和计算对象的粒子数 N 有关)的四次方成正比。因此一直到 60 年代以后，随着计算机

的发展才很快发展起来。即使是在今天，实际上很难用从头计算法来计算包含成千上万个电子的药物和生物分子，对于这类分子(尤其是像蛋白质和核酸这样的多聚物)，还有待提出一些新的思想和建立一些新的方法。

但是,对于 STO≥50 的一些“大分子”（这里所指当然不包括通常所指的生物大分子，如蛋白质、核酸等)的从头计算技术正日益发展。陈凯先等[22]对提高大分子从头计算速度的技术，作了较深入的研究,并成功地编制了大分子从头计算的通用程序 MQM 81 和模型势方法计算程序 MQMMP。

3. 从头计算法的发展

尽管,从头计算法是目前为止最严格和最精确的分子轨道法。然而,正如前面所述，由于它的计算量庞大,占用计算机内存多,耗费机时长,故而使人“望而生畏”,尤其是距离实现生物大分子的从头计算还有相当长的路程。因而,对于量子生物学的理论计算而言,在一定程度上从头计算法还只能是纸上谈兵。即便如此，人们仍旧在不断地寻求和探索这一理论计算方法的改进和发展。在这些工作中值得指出的有：赝势价电子从头计算法;浮动球 Gauss 轨道法;从头算碎片间的相互作用加和法。

(1) 赝势价电子从头计算法[23,24]

赝势(pseudopotential)又称模型势(model potential)价电子从头计算法,是近十多年来迅速发展起来的一种重要的量子化学计算方法。主要用于较大的分子，特别是含重原子的化合物和原子簇。

赝势法也是一种不依赖于任何经验参数的非经验方法，它与全电子从头计算法的差别在于它把变分和自洽场处理仅用于价电子,对于内层电子的影响和贡献则通过一组有效势(effective potential) 算子来模拟。这组有效算子是在原子状态下通过拟合全电子从头计算的结果来确定的,进而直接用于分子体系。可见,赝势法的计算量比从头计算要小得多。譬如对氨分子的计算,占用机时约为从头计算法的 1/10。如果计算的分子含内层电子愈多，则节省的机时也愈多，甚至可降至 1/100。

大量的计算实践表明,赝势法的计算结果可相当好地逼近全电子从头计算法。迄今为止,已经提出了多种赝势法计算方案，并在配合物、过渡金属、有机化合物、原子簇化合物和药物分子的计算上取得了成功。我国的科学工作者也对赝势法进行了深入的研究,如陈凯先、金络歆等编制的通用程序在若干研究课题中获得了成功。

(2) 浮动球 Gauss 轨道方法[17]

浮动球 Gauss 轨道法(floating spherical gaussian orbital method，简称 FSGO) 是一种简单的从头计算方法。它主要用于讨论具有偶数个电子体系的闭壳层的基态。该方法运用定位分子轨道模型,并假定每一个定位分子轨道用一个完全球对称的 Gauss 函数表述,且每一对电子占据一个这样的轨道。再由这些占据轨道直接构成单 Slater 行列式波函数来表示体系的状态，进而采用非线性变分法求解 Schrödinger 方程。

在量子生物学研究中,Christofferson等提出的“分子碎片法”(molecular fragment method)获得了较多的应用。由于该方法采用浮动球 Gauss 轨道,因此又称为分子碎片浮动球 Gauss 轨

道方法（简称 MF-FSGO）。这一方法在研究多环芳烃(PAH)致癌机理中，对若干反应中间物(如其中的 9-甲基鸟嘌呤与三羟基正碳离子苯并[a]芘的复合物)作了量子化学计算，效果较为满意[25]。

(3) 从头算碎片间相互作用加和法

从头算碎片间的相互作用加和法（sum of interactions between fragments computer ab initio 法，简称 SIBFA），是近些年由 Pullman 等人[26,27]建立的计算生物大分子的量子化学方法。由于生物大分子体系的庞大，在现今整体从头计算尚无法实现的情况下，需要发展可行的有针对性的方法。于是，SIBFA 应运而生。它是以局部结构(Fragment)的从头计算为基础，采用一些经验和半经验性的模型，计算比较复杂的庞大体系。

已经发表的方法有 SIBFA1 和SIBFA 2。其中，SIBFA 1 系由于分子构象的改变而引起分子内部能量的变化，通过碎片之间相互作用加和的可变部分获得，用下式表示：

$$\Delta E_{\text{conf}} = E_{\text{mtp}} + E_{\text{pol}} + E_{\text{rep}} + E_{\text{disp}} + E_{\text{tor}} \tag{2,130}$$

对于 SIBFA 2，它适用于 Drug-tetranucleotide 分子间的相互作用能计算，其表达式为

$$\Delta E_{\text{inter}} = E_{\text{mtp}} + E_{\text{pol}} + E_{\text{disp}} + E_{\text{ct}} \tag{2.131}$$

在(2.130)和(2.131)式中，E_{mtp}和E_{pol}分别表示静电的和极化的能量贡献，它们可通过碎片的自洽场从头计算所得的波函数经多极展开而获得，E_{rep}和 E_{disp}分别是排斥和色散的能量部分，E_{tor}是扭转能的贡献，E_{ct}是电荷转移能项。由以上两式可见，这两种方法分别考虑了分子内和分子间相互作用能，在理论和实践上都可能有广阔的应用前景。

陈凯先等[28]应用此法计算了抗癌药物 Mitoxantrone 与 DNA 的相互作用，初步显示出这种方法用于大分子计算的潜力和作用。

2.2.4 X_α 方 法

从头计算法计算工作量太大，实际上仍无法对生物大分子进行计算，而半经验方法又过于近似，往往得不到令人满意的计算结果。介于二者之间的有所谓 X_α方法。

1. 计算原理

X_α方法是 Slater[29]在 1951 年首先提出来的，他将电子交换作用能用其统计平均来近似，代入 Hartree-Fock 方程，得到 X_α 方程。

这种方法已成为计算含重原子的一些分子的重要工具，并广泛地应用在化学、生物学、固体物理和表面科学等学科。

(1) 能量表达式

$$E_{X_\alpha} = \sum_i n_i \int u_i^*(1) f_1 u_i(1) \mathrm{d}\tau_1 + \frac{1}{2}\int \rho(1) U_c(1) \mathrm{d}\tau_1 + \frac{1}{2}\int \rho(1) U_{X_\alpha}(1) \mathrm{d}\tau_1 \tag{2.132}$$

式中 u_i 为包含自旋的单粒子波函数(分子轨道)，n_i为占据 u_i轨道的电子数(0 或 1)，f_1为单粒子算符，$\rho(1)$为电荷密度，$U_c(1)$为库仑作用能，$U_{X_\alpha}(1)$为交换作用能，它们的表达式如下：

$$
\left.\begin{aligned}
f_1 &= -\frac{1}{2}\nabla_1^2 - \sum_p \frac{Z_p}{r_{ip}} \\
\rho(1) &= \sum_i n_i u_i(1) u_i^*(1) \\
U_c(1) &= \int \frac{\rho(2)}{r_{12}} \mathrm{d}\tau_2 \\
U_{X\alpha}(1) &= -\frac{9}{2}\alpha\left(\frac{3}{8\pi}\rho(1)\right)^{\frac{1}{3}}
\end{aligned}\right\} \tag{2.133}
$$

(2.132)式中的第一项，代表电子的动能及其与核电荷的吸引能，第二项代表电子间的排斥能，这两项和从头计算的表达式一样，除了三个基本近似外，没有引进其它近似。第三项是交换作用能，在采用了统计平均近似后，和 $\rho^{\frac{1}{3}}$ 成比例。参数 α 则是在大量工作的基础上引入的。

(2) X_α 方程

将(2.132)式对 u_i 变分，求能量的极小值，利用分子轨道 u_i 的正交归一化性质，就得到 X_α 方程

$$
fu_i + U_c u_i + \frac{2}{3} U_{X\alpha} u_i = \varepsilon_i u_i \tag{2.134}
$$

系数 $\frac{2}{3}$ 是对 u_i 微分而得，一般文献中的 X_α 方程为

$$
fu_i + V_C u_i + V_{X\alpha} u_i = \varepsilon_i u_i \tag{2.135}
$$

式中

$$
V_c = U_c; V_{X\alpha} = \frac{2}{3} U_{X\alpha} = -3\alpha\left(\frac{3}{8\pi}\rho\right)^{\frac{1}{3}} \tag{2.136}
$$

(2.135)是自旋限制的 X_α 方程。

(3) X_α 方程的性质

X_α 方程严格满足 Fermi-Dirac 统计

$$
\varepsilon_i = \frac{\partial E_{X\alpha}}{\partial n_i} \tag{2.137}
$$

分子轨道的能量等于总能量对该轨道电子占有数的偏微分。

X_α 方程校正了 Koopman 定理中包含的“冻结”近似，引入了“过渡态”，电离能等于过渡态分子轨道的能量

$$
-I_k = \varepsilon_k\left(n_k = \frac{1}{2}\right) \tag{2.138}
$$

ε_k 表示半个电子拿走，半个电子留在 u_k 时的分子轨道的能量。

X_α 方程满足 Viril 定理

$$
2T + V = 0
$$

动能 T 与位能 V 之比等于负 1/2。

关于参数 α 值的确定，有三种方法，运用三种方法所得结果基本相同。周期表中从 $Z=1$ 到 86 的原子的优选 α 值可参阅文献[30]。

2. 多重散射 X_α 方法

原子具有球对称性,以及晶体具有空间周期性时,X_α 方程易于求解。把 X_α 方程用来计算分子结构是在 Johnson[31]将讨论固体能带的多重散射技术改用于计算多原子分子的束缚单电子本征态的基础上发展起来的。

该方法能给出 Hartree-Fock 原型方程的较真实解而无需过于繁杂的计算, 这对于处理较大分子尤为有利。原则上,它还是一种从头计算法,但通过引入参数 α 又具备了半经验方法的灵活性。它不必把分子实和价电子(或 σ 电子和 π 电子)分开处理,也没有忽略它们的相互作用,各部分均在同一计算中自洽解决。

多重散射 X_α 方法采用"松饼罐头"(Muffin-tin)近似,把整个分子视为一个大球,每个原子视为一个小球。于是,整个分子体系被分为三个区域(图 2-5),即原子内区域 I, 原子球间区域 II 及分子球外区域 III。在每个区域,对分子轨道间的交换作用能都采用 Slater 的统计平均近似,引入参数因子 α。 把原子内区域 I 和分子球外区域 III 的势能对方向求平均,近似为只是 r 的函数,这些区域的 X_α 方程的解具有中心力场近似下 Schrödinger 方程的解的形式。把原子球间区域 II 的势能对体积求平均,近似作为常数,该区域的 X_α 方程的解是球 Bessl 方程的解的形式,分为两部分,一部分为入射波,一部分为散射波,各部分均由 L 个(球谐函数基的数目)分波线组合而成。

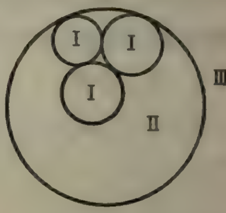


图 2-5 分子区域划分示意图

根据多重散射理论, 三个区域的波函数及其一阶导数在边界上应是连续的。于是可以导出一系列迅速收敛的久期方程, 求解这些久期方程而得到分子轨道能量和波函数。整个求解过程仍是一个叠代自洽过程:先选择一套试探分子轨道,计算电荷密度 $\rho(1)$ 而得各区域的势函数,随后求解各区域的 X_α 方程,得到组合系数待定的各区域的普遍解。再通过求解前述久期方程及波函数系数的代数方程组而得到能级和新的波函数, 由新的波函数去计算电荷密度和新的势能进行下一次叠代,直到自洽为止。对 X_α 方法的详细讨论可参阅文献[31—33]。

目前,用 X_α 方法计算分子结构的工作已逐渐多起来了,一般结果还是比较好的。近年来,Karplus 等应用 X_α 方法计算了生物分子[34],有关这一方法在生物学中的应用正在进行中。

2.2.5 分子静电势法

分子静电势(molecular electrostatic potentials)是分子体系的一项重要性质。在量子化学中,通常的 Mulliken 集居数分析方法在计算各原子的净电荷时,并不考虑各自在电负性上的差别,而是将两个原子间的电子重叠集居数平分给这两个原子。研究表明,这一浅略的处理有时可能会导致物理图象的严重歪曲,因此是不够合理的。Kaufman 和 Pullman 等建立了分子静电势方法,已经成为量子生物学计算中的一个重要组成部分[35-37]。

分子静电势的定义是:空间某点的静电势表示电场与位于该点的单位正电荷之间的相互作用能。

$$V(\vec{r})=\sum_{A}\frac{Z_A}{|R_A-\vec{r}|}-\int\frac{\rho(\vec{r}')\mathrm{d}\vec{r}'}{|\vec{r}'-\vec{r}|} \tag{2.139}$$

上式两项具有明确的物理意义，前一项表示原子核对该点的静电势，后一项表示电子对该点静电势的贡献。

从分子静电势的定义出发，可以看到它的物理意义与经典物理学中的含义是一致的。分子静电势的计算以及据此绘制的分子静电势图，可提供在 Hartree-Fock 近似下分子静电性质的准确、客观和直观的图象，从而避免 Mulliken 方法的缺陷。

分子静电势法在药物分子，核酸和蛋白质等生物分子的计算中获得了广泛的应用。按照静电势图，我们可以比较分子的正、负电势区域，以此来判断分子间相互作用的可能的活性中心，并可对亲电进攻的反应性作出可能的分析。分子静电势法为药物一受体相互作用机制的探讨，酶活性部位的分析，提供了一种非常有用的工具。此外，亦可借助对同源化合物分子静电势的计算，并将活性部位的负电势值与生物活性进行相关分析计算，从而建立定量构效关系 (QSAR)方程式。

2.2.5.1　从头计算法的分子静电势计算[38]

从头计算法得到的波函数是迄今最精确的波函数，而建立在从头计算法基础上的分子静电势计算已成为目前应用最广泛也是最精确的分子静电势法。

根据分子静电势的定义，空间某一点 P 的静电势可以写成：

$$V(\mathrm{P})=\sum_{a}\frac{Z_a}{r_{aP}}-\left\langle \Psi \left| \sum_{i} V(i) \right| \Psi \right\rangle \tag{2.140}$$

式中的两项分别表示原子核和电子在 P 点产生的电势，a、i 标志各个原子核和电子，Z 是核电荷数，$V(i)$ 是算子$\frac{1}{r_{Pi}}$。

运用密度矩阵理论知识，上式后一项可推导为

$$\begin{aligned}
&\left\langle \Psi \left| \sum_{i} V(i) \right| \Psi \right\rangle \\
&= N\langle \Psi | V(1) | \Psi \rangle \\
&= N\int V(1)\mathrm{d}\tau_1 \int \Psi^* \Psi \mathrm{d}\tau_2 \cdots\cdots \mathrm{d}\tau_N \\
&= \int V(1)\cdot \rho_1(\vec{X}_1)\mathrm{d}\tau_1 \\
&= T_r V(1)\cdot \rho_1
\end{aligned} \tag{2.141}$$

式中 ρ_1 是一级约化密度矩阵

$$\rho_1 = 2\sum_{i,j} \Phi_j(\vec{r}_1)\cdot \boldsymbol{R}_{ji}\cdot \Phi_i(\vec{r}_1')^* \tag{2.142}$$

把(2.141)和(2.142)代入(2.140)，即可得到静电势计算的普遍公式

$$V(P)=\sum_{a}\frac{Z_a}{r_{aP}}-2\sum_{i,j}\langle \Phi_i | V | \Phi_j \rangle \cdot \boldsymbol{R}_{ji} \tag{2.143}$$

这里，Φ_i、Φ_j 是组成分子轨道的基函数，$\boldsymbol{R}_{ji}$ 是分子体系的密度矩阵元。

$$\langle\Phi_i|V|\Phi_j\rangle=\int\frac{\Phi_i^*(r)\cdot\Phi_j(r)}{r_p}d\tau \qquad (2.144)$$

在从头计算法中，当采用 GTO 基函数时，(2.144)式的值可由下式来计算：

$$\langle\Phi_i|V|\Phi_j\rangle$$

$$=N_A\cdot N_B\cdot\frac{2\pi}{\beta}e^{-\alpha_1\alpha_2\frac{\overline{AB}^2}{\beta}}$$

$$\cdot\sum_{l,r,u}\cdot\sum_{m,s,v}\cdot\sum_{n,t,w}A_{lru}\cdot A_{msv}\cdot A_{ntw}\cdot F_\nu\left(\frac{\overline{PQ^2}}{4\delta}\right) \qquad (2.145)$$

式(2.145)中，A、B 分别为两个 GTO 的中心，N_A、N_B 分别为归一化系数，Q 为 $x_i\ x_j$ 合併成单一中心后的中心点。

$$A_{lru}=(-1)^l\cdot l!\frac{(-1)^u\cdot\delta^{r+u}\cdot\overline{PQ_x}^{\,l-2r-2u}}{r!u!(l-2r-2u)!}f_l(l_1,l_2,\overline{QA_x},\ \overline{QB_x})$$

$$\beta=\alpha_1+\alpha_2;\quad \delta=\frac{1}{4\beta};$$

$$\nu=(l+m+n)-2(r+s+t)-(u+v+w)$$

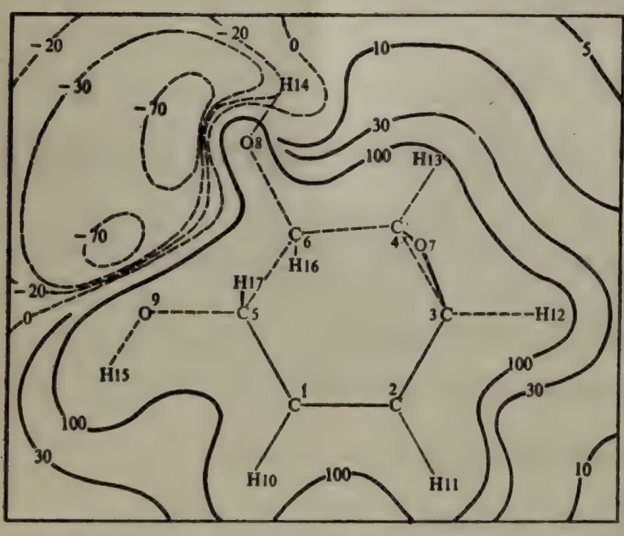


图 2-6 反式($eq\cdot eq'$)异构体，分子静电势图

$$F_\nu\left(\frac{\overline{PQ^2}}{4\delta}\right)=\int_0^1 S^{2\nu}\cdot e^{-\frac{\overline{PQ^2}\cdot S^2}{4\delta}}dS$$

式中的求和范围是

$$l:\ 0\longrightarrow l_1+l_2$$

$$r:\ 0\longrightarrow\left[\frac{l}{2}\right]$$

$$u:\ 0\longrightarrow\left[\frac{l-2r}{2}\right]$$

在上述数学公式的基础上，即可设计程序进行计算。文献[38]报导了用 FORTRAN IV 语言编制的分子静电势计算程序 ESP，可准确方便地计算分子体系所在空间任一平面上的静电势；也可由空间曲面方程的控制，计算某一曲面上的分子静电势。它所需的波函数，可由程序自动读入事先经从头算自洽后存入计算机磁盘的数值。该程序适用于计算生物大分子。

文献[39]报导了应用分子静电势计算程序 ESP 计算的二醇环氧化物的模拟分子的三种空间构型的分子静电势，图 2-6 至图 2-8 显示了这三种构型的静电势图。

在以上三图中，图所在的平面是 xy 平面，z 轴与这个平面垂直。图中虚线表示负电势等值线，实线代表正电势等值线。图 2-6 和图 2-8 非常相似，而它们与图 2-7 显然不同。这表明在两个平伏键型(eq,eq')异构体周围空间，分子静电场具有基本相同的特征，而直立键型(ax,ax')异构体则与它们有很大差别。对于平伏键型异构体，分子静电势的负区域位于 xy 平面以及接

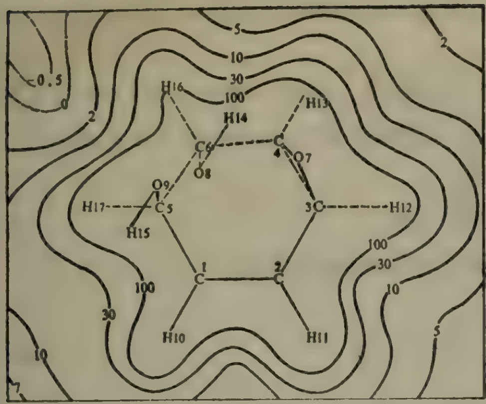


图 2-7　顺式(ax,ax')异构体：分子静电势图

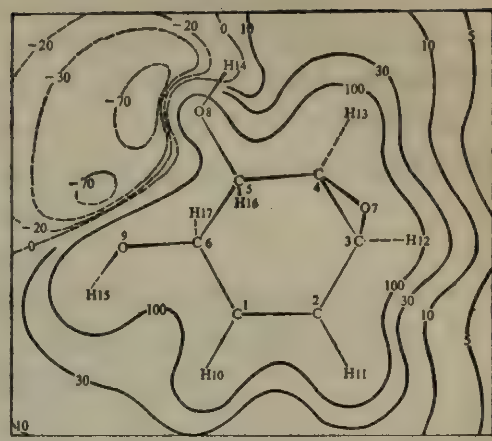
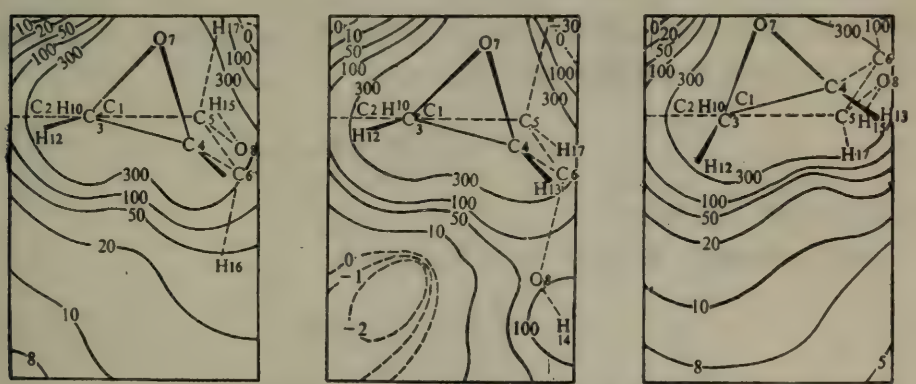


图 2-8　顺式(eq,eq')异构体：分子静电势图



I　反式(eq,eq')　　II　cis(ax,ax')　　III　cis(eq,eq')

图 2-9　垂直平面上的分子静电势

近这一平面的空间；而对于直立键型异构体，这些负区域主要地是分布在 xy 平面的上方和下方。这样，对于前者而言，从分子静电势的负区域到 C_3 原子的距离就比后者远。后者的这一特征不利于亲核部位从 C_3 原子背侧的进攻。进一步对通过 $C_3—C_4$ 的联结线且和 xy 平面垂直的平面上的分子静电势的计算，绘成了图 2-9(I—III)。图中所示的平面垂直于 xy 平面，纵轴是 z 方向。在这样的坐标下，亲核部位从环氧背侧进攻 C_3 的路线，就是从图的下方向上指向 C_3 原子。这就更清楚地表明直立键型异构体不利于亲核部位从 C_3 原子背侧进攻。图中显示对直立键型，静电势图的下方有一负区域阻挡着向 C_3 原子亲电进攻的路径，使得这种进攻大为困难。而对于两个平伏键型异构体，完全不存在这种困难，这就对多环芳烃的最终致癌物——二醇环氧化物平伏

键型异构体致癌活性明显高于直立键型异构体这一实验事实给出了一个直观有说服力的解释。

2.2.5.2 CNDO 法分子静电势计算[40]

一般而论，分子静电势的获得是依赖于从头计算所得到的波函数。但是，由于这种非经验的量子化学计算方法在处理药物分子特别是生物大分子时，确实存在着很大的困难。因而，Pullman 等又建立了用 CNDO/2 法获得的波函数来计算分子静电势的方法。文献[40]报道了 CNDO/2 法分子静电势计算公式的推导及程序化。

在 CNDO/2 近似下，公式(2.140)中，Z_a 取价电子数目，而后一项积分在 CNDO 法采用的零微分重叠(ZDO)近似下，$\langle \Phi_i | V | \Phi_j \rangle$ 只有当 $i=j$，并同属于同一原子的条件下才需计算

$$\langle \Phi_\mu^A | \hat{V}_P | \Phi_\nu^B \rangle = \delta_{AB}\delta_{\mu\nu}\langle \Phi_\mu^A | \hat{V}_P | \Phi_\mu^A \rangle \tag{2.146}$$

可采用多种方案处理积分 $\langle \Phi_\mu^A | \hat{V}_P | \Phi_\mu^A \rangle$ 的计算，以下介绍两种近似：

1. 近似 1：采用 CNDO/2 方法，即

$$\langle \Phi_\mu^A | \hat{V}_P | \Phi_\mu^A \rangle = r_{AP} \cdot Z_P \tag{2.147}$$

其中

$$r_{AP} = \iint S_A^2(1) \frac{1}{r_{12}} S_P^2(2) \mathrm{d}\tau_1 \mathrm{d}\tau_2$$

有关 r_{AP} 的计算已在 CNDO/2 法中论述。在实际计算中，由于 P 点的静电势事实上等于位于该点的单位正电荷与电场的相互作用能，所以上式中 S_P 即可用氢原子的 1 s 波函数代入，并且 $Z_P=1$，于是静电势计算公式可简化为：

$$V(P) = \sum_a \left(\frac{Z_a}{r_{PA}} - P_{AA} r_{AB} \right) \tag{2.148}$$

式中的 P_{AA} 为密度矩阵对角元。

2. 近似 2：精确地计算 $\langle \Phi_\mu^A | \hat{V}_P | \Phi_\mu^A \rangle$，即计算积分

$$\int (\Phi_\mu^A(1))^2 \frac{1}{r_{P1}} \mathrm{d}\tau$$

按照 CNDO 法的原则，取 Φ_μ^A 为 s 型轨道，将 $\frac{1}{r_{P1}}$ 作极坐标下单中心展开，并利用Legendre 多项式的正交性，导得下列计算公式

$$\begin{aligned}
&\langle \Phi_\mu^A | \hat{V}_P | \Phi_\mu^A \rangle \\
&= \int \chi^2(r,\theta,P) \frac{1}{r_{P1}} \mathrm{d}\tau \\
&= -\frac{(2\zeta)^{2n+1}}{(2n)!} \cdot e^{-2\zeta P} \left[\sum_{i=1}^{2n} \frac{1}{(2\zeta)^i} \cdot \frac{(2n-1)!(i-1)}{(2n-i+1)!} \right. \\
&\quad \left. \cdot P^{2n-i} \right] - \frac{1}{P}(e^{-2\zeta P} - 1)
\end{aligned} \tag{2.149}$$

式中 n 和 ζ 分别是 Slater 轨道的主量子数和轨道指数，P 为 A 原子到空间一点 P 的距离。在设

性计程序时.(2.149)式进一步按 n 的不同数值展开成多项式加和的形式。n 从 1—3 的展开式为

$$n=1 \quad -\xi\cdot e^{-2\zeta P}-\frac{1}{P}(e^{-2\zeta P}-1) \tag{2.150}$$

$$\begin{aligned} n=2 \quad & -e^{-2\zeta P}\cdot\left[\frac{1}{24}(2\zeta)^3\cdot P^2+\frac{1}{4}(2\zeta)^2\cdot P+\frac{3}{4}(2\zeta)\right] \\ & -\frac{1}{P}(e^{-2\zeta P}-1) \end{aligned} \tag{2.151}$$

$$\begin{aligned} n=3 \quad & -e^{-2\zeta P}\cdot\left[(2\zeta)^5\cdot\frac{1}{720}P^4+\frac{1}{72}(2\zeta)^4\cdot P^3+\frac{1}{12}(2\zeta)^3\cdot P^2\right. \\ & \left.+\frac{1}{3}(2\zeta)^2\cdot P+\frac{5}{6}(2\zeta)\right]-\frac{1}{P}(e^{-2\zeta P}-1) \end{aligned} \tag{2.152}$$

用上述公式计算积分后静电势可由下式获得

$$V(P)=\sum_{a}\frac{Z_a}{r_{PA}}-\sum_{\mu\mu}P_{\mu\mu}\cdot\langle\Phi_\mu^A|\hat{V}_P|\Phi_\mu^A\rangle \tag{2.153}$$

根据以上两种近似方案,文献[40]给出了使用两种公式计算分子静电势的程序 CMEP。用这一程序对多环芳烃致癌机理湾区历程中最终致癌物模拟分子之一——反式平伏键型[*trans* (*eq*,*eq'*)]分子进行了计算,图 2-10,2-11 给出了两种近似的 xy 平面的分子静电势图。

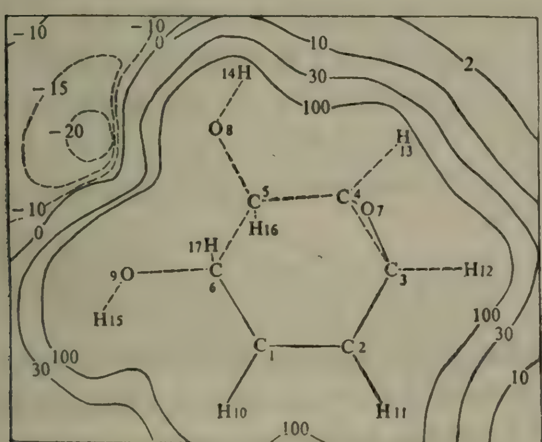


图 2-10 CNDO/2 法近似 1 静电势图

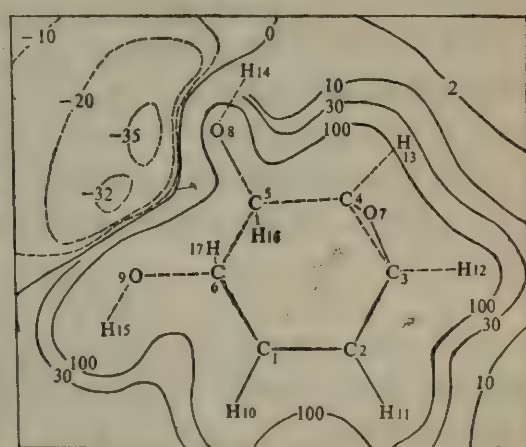


图 2-11 CODO/2 法近似 2 静电势图

将这些图与从头计算(STO-3 G)的静电势图(图 2-6)比较,可以看出它们的图形相近,正负区域明显可分。由近似 1 获得的图 2-9 其静电势凸凹线比从头算浅;而近似 2 重合得较好,差别仅仅在于所获得最大负电势的绝对值约为从头计算值的 1/3—1/2, 与 Pullman 报导的结果[41]一致。这是由于 CNDO 法数学模型较粗糙所致。尽管如此,由于它的图形与从头算吻合较好,可满足同类型化合物相对比较研究的需要,且易应用于大分子计算。因此,在量子生物学研究中已成为计算分子静电势的一种便利工具[42-46]。

此外，我们可用赝势价电子从头计算法获得的波函数来计算分子静电势。文献[47]报道了赝势价电子从头算静电势计算程序ESP-2,并已成功地应用于顺、反式二氯二氨合铂$[Pt(NH_3)_2]Cl_2$ 和$[Pt(NH_3)_2]^{+2}$ 阳离子的分子静电势计算。

使用由其它半经验分子轨道法获得的波函数进行分子静电势的计算的报道也很多，通常在生物学和药物学中应用的有 INDO 法[48],MINDO/3 法[49],及近年来报道的 MNDO 法[14]。由于这些半经验方法都基于零微分重叠(ZDO)原理，因此它们计算分子静电势的方法一般都基于(2.153)式。

2.2.5.3 生物分子的分子静电势计算

由于生物分子(包括药物分子)是一些较大的分子体系，因而使用量子化学方法获得波函数较为困难。

但是,从近年来在量子生物中有关分子静电势的研究论文看，使用从头计算法(包括赝势价电子从头算法)计算静电势的约占2/3。纵观这些文献，发现它们都采用了计算分子片断或计算模拟分子的办法，来逾越由于分子体系太大而带来无法计算的障碍。在研究系列化合物的作用机制时,往往计算这些分子主要作用部位的分子片断,而将侧链或相同部位略去,用氢原子饱和。这种片断计算也可获得令人满意的结果[50-52],模拟分子的计算也是基于上述原则，在单个分子构型变化的计算中常常采用此法[39]。由于它们是基于从头算非经验的量子化学计算的水平，因而这些分子静电势计算结果优于半经验法（如 CNDO/2）对整个分子的计算结果。甚至有的 CNDO/2 法计算静电势也采用了这种分子片断的近似计算,也获得了良好的结果[44,45]。

一般药物均包含芳香环,早期的分子静电势计算常以芳环为平面，直接计算 xy 平面($z=0$)的分子静电势。后来发现当 $z\geqslant1.75$,平行于 xy 平面上计算静电势，可以较好地考虑 π 电子的静电作用,获得了很好的结果[50-52],这种处理方法已逐渐被大家所接受。

我国计算分子静电势的工作起步较早，其研究工作几乎与生物分子从头计算程序同步发表[38],可以相信,分子静电势的应用将会越来越广泛。

2.2.6 PCILO 计算法

PCILO 法亦即微扰组态相互作用定域分子轨道法 (perturbation configuration interaction using localized orbitals)，是由 Diner 和 Malrieu 等于 1967 年起开始建立的半经验的量子化学分子轨道法[53-56]。这个方法是以全价电子半经验的量子化学方法为起点，计算分子整个构象的一种简便方法。由于 PCILO 法考虑到组态相互作用,在相当程度上弥补了传统的自洽场轨道所忽略的“相关能”,因而计算结果较好。该法还运用微扰理论,直接由零级、二级及三级能各项修正值加合而得,使能量计算避免了自洽叠代的过程,致使计算速度远比一般的半经验分子轨道法为快。该方法还采用了“二中心——双电子”化学键及孤对电子、空轨道来组建分子的定域轨道。其零级近似对应 Lewis 结构式,具有化学的直观性,在量子生物学中的应用尤具特色。

Pullman 认为 PCILO 计算法是研究核酸及生物分子的有效方法[57]。Richards 也在其第二版《量子药理学》一书中认为,PCILO 法是药理学领域中最广泛使用的分子轨道方法之一[58]。他

在小结应用于量子药理学计算的分子轨道方法时，给 PCILO 法以显著的地位，认为它的计算结果在某种程度上优于IEHT(叠代的扩充Hückel 分子轨道法)，与CNDO/2 结果相近(图 2-12)。

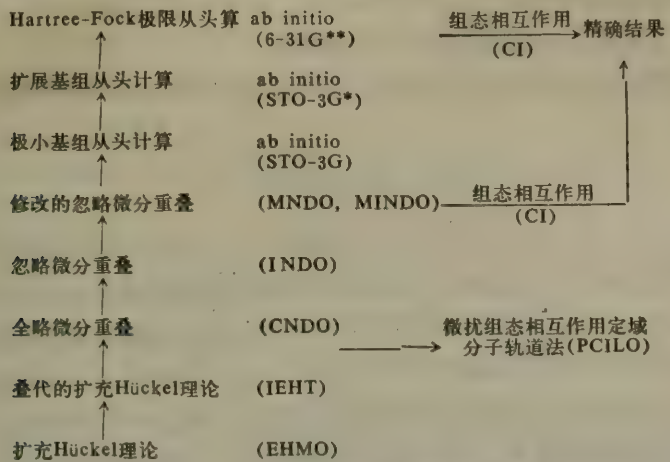


图 2-12　各种不同近似的分子轨道法比较

2.2.6.1　PCILO 法基本原理

1. 组态相互作用 (cofiguration interaction——CI)

在分子轨道理论中，当用自洽场的方法(SCF)求解 Hartree-Fock 方程时，依照 Pauli 原理，将具有不同自旋(自旋 α 和自旋 β)的两个电子置于同一轨道，即允许同一地点找到两个电子，这就与电子相关理论发生了冲突。电子相关理论认为，具有相同电荷的粒子(如电子)由于Coulomb 排斥作用而彼此尽量避开。

由于经典的 SCF 方法存在的所谓相关误差，因而在分子轨道理论中，常常使用组态相互作用(CI)的方法来修正相关误差。

在求解 SCF-LCAO 方程时，可得到组成行列式波函数 Φ_K 的分子轨道，它分为占据轨道和空轨道。在 CI 处理中，Φ_K 的开头的一些组分，适当地用空轨道代替，以使这些行列式 Φ_K 相应于该体系的激发组态，这意味着一个电子被激发到较高能量的轨道上。组态可以是单重 (singly)、双重(doubly)和三重(triply)的。图 2-13 表示了各个组态的分子轨道能级图。组态 A，称为闭

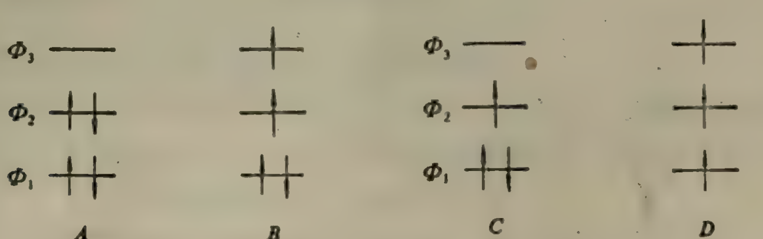


图 2-13　组态示意图

壳层组态，它对应的波函数为单重态。组态C，含有单电子的占据轨道，称为开壳层组态。由于Φ_2中的一个电子可以取α和β自旋态中的任一个，因此C对应一个双重态。组态B具有二个开壳层的电子，其电子自旋的组合对应一个单重态和一个三重态。组态D有三个不成对电子，按照各自的自旋组合可以组成一个四重态和两个双重态。

显然，依照一种适当的方式混合这些组态，可使真实的波函数成为所有可能组态的线性叠加，亦即

$$\overline{\Psi}=\sum_K C_K\Phi_K \tag{2.154}$$

式中每一个Φ_K代表一个Slater行列式，即一个组态，C_K为该组态的权重系数，可由变分法求得。这样的波函数就把电子相关包括进去了，用足够数目的不同的Φ_K，可以使$\overline{\Psi}$逼近真实波函数，以达到预期的要求。当然，采用组态相互作用，就要求解很大的组态作用矩阵，以便获得系数C_K。看起来，这样做似乎带来了很大的困难，但事实上，对基态而言，只有和基态能量相近的态才有较大的贡献，对于激发态，只有具有相同及相近激发度的组态才有较大的贡献。这样，只需采用有限个组态，就能得到较好的近似。因此CI常常合并于各种近似方法，如P. P. P.法，CNDO/2法，ab initio法等中，而PCILO法则是以CI为基点的方法。

2．微扰组态相互作用(perturbation configuration interaction——PCI)

为寻找一个简便而又准确地求解CI问题的办法，除变分法外，还可以采用微扰技术。即把Φ_K看作是未微扰的Hamiltonian H_0的本征向量，它与真实的Hamiltonian $\mathscr{H}$(已包括相关作用)仅差一个小的微扰项λV

$$\mathscr{H}=H_0+\lambda V \tag{2.155}$$

H_0的对角矩阵元为

$$(H_0)_{kk}=E_k \tag{2.156}$$

则可以在$\lambda=0$处对真实解的能量和波函数作Taylor展开

$$\varepsilon=E_K+\lambda E_1'+\lambda^2E_2'+\cdots\cdots \tag{2.157}$$

$$\overline{\Psi}=\Phi_K+\lambda\Psi_1'+\lambda^2\Psi_2'+\cdots\cdots \tag{2.158}$$

当K取0时，E_i'和Ψ_i'对E_0和Φ_0的一级、二级、三级等的修正项可利用Rayleigh-Schrödinger微扰理论经典公式，依据V的矩阵元和H_0的本征值获得

$$E_1'=V_{00} \tag{2.159}$$

$$E_2'=\sum_{K\neq0}\frac{|V_{0K}|^2}{E_0-E_K} \tag{2.160}$$

$$E_3'=\sum_{K\neq0}\sum_{l}\frac{V_{K0}V_{l0}V_{0K}}{(E_0-E_K)(E_0-E_l)} \tag{2.161}$$

在PCI计算中，要求

$$\mathscr{H}_{IJ}=V_{IJ}\qquad I\neq J \tag{2.162}$$

$$\mathscr{H}_{II}=(H_0)_{II}\quad 对于全部 I \tag{2.163}$$

这里$\mathscr{H}$已是真实的Hamiltonian，如此，E_K是用$\mathscr{H}$计算的单行列式能量，所有$V_{II}=0$，于是一级修正$E_1'=0$，而且V的其他必要的矩阵元是用$\mathscr{H}$计算的CI的一般矩阵元。这样，就容易求得结

果了。

3．定域轨道(localized orbitals)

通常求解 Fock 矩阵所得的分子轨道，无论其属于 σ 或 π 表示，都非定域于整个分子，因而，这种分子轨道是对称适应的，即它们形成由分子对称性所对应的点群的不可约表示的基。另一方面，定域分子轨道正等价于分子成键 Lewis 结构式。如 CH_4 分子的定域分子轨道的产生，其方向必须各自沿一个 C—H 键。

在 PCILO 法中正是由于采取了定域分子轨道 $\overline{\Phi}$，因此通过 U 矩阵变换可将非定域分子轨道变成定域分子轨道 $\overline{\Psi}$。

$$\overline{\Psi}=\overline{\Phi}U \tag{2.164}$$

2.2.6.2 PCILO 法的一般概要

PCILO 法一般可概括为以下步骤：

1．分子轨道的构成

PCILO 法的独到之处完全在于定域轨道的选择。首先用 Del Re 方法构成杂化轨道。杂化轨道的取向由成键原子或虚拟原子(或称哑原子)确定，通过满足最大重叠判据及正交化而得。亦即起始分子轨道根据化学式提供的图形被构造成成键轨道和孤对轨道，适当的原子杂化轨道两两结合：

$$i=\alpha\chi_1+\beta\chi_2 \quad (成键) \tag{2.165}$$

相应于每个成键轨道构筑反键轨道

$$i^*=-\beta\chi_1+\alpha\chi_2 \quad (反键) \tag{2.166}$$

(对于孤对电子，χ_2 对应于无限大被舍掉；β 和 i^* 等于零)。

由此可见，以上这种选择完全适用于从头计算方法。然而，通常所指的 PCILO 法是指引进全价电子的半经验方法。

首先，它利用完全的 ZDO(零微分重叠)假设。其次，它采用 CNDO/2 法计算 Coulomb 积分和 Core 矩阵元的积分近似。并且，按照 CNDO/2 近似

$$\alpha^2+\beta^2=1 \tag{2.167}$$

2．成键轨道用于构成 Slater 行列式

利用 Del Re 方法，由通常的 Slater 原子轨道构造适当的杂化轨道，以便保证在化学键上的最大重叠以及键间的正交性。ZDO 条件对起始成键轨道集合自动正交有明显的好处。因为零级基态行列式在全部 i 轨道上建立起来，它就是分子的零级波函数。

$$\Phi_0=|i_1\,i_2\cdots\cdots i_K\cdots i_N| \tag{2.168}$$

3．用反键轨道构成相当于激发组态的 Slater 行列式：

利用反键轨道 i^* 来构造所有可能的“激发”的 Slater 行列式，这样，总的 Hamiltonian $\mathscr{H}$，就可完全确定 E_K 和必需的微扰矩阵元 V_{IK}。故而，对能量的第二级相关包括

(1) 单激发 $i\longrightarrow i^*$(极化项)或 $i\longrightarrow j^*$(离域项)

(2) 利用 ZDO 双键激发简化到

$$ii\longrightarrow i^*i^* \quad (键内'相关')$$

$$ij \longrightarrow i^*j^* \quad (\text{键间'相关'})$$

而第三级相关项(2.161式)涉及前面的组态间的相互作用，亦即

——单激发与单激发；

——单激发与双激发；

——双激发与双激发；

第一个结合是由极化和(或)非定域相关构成，第二和第三结合涉及再“相关”。

以上清楚地表明 PCILO 法根据具有和不具有从一个键到另一个键的电子转移激发作用而分类，这样分类较充分地体现了组态激发中所包含的键特征性。

这样，波函数即可对应为

$$\Phi_I = |i_1 i_2 \cdots\cdots i_K^* \cdots\cdots i_N| \quad \text{单激发态} \qquad (2.169)$$

$$\Phi_I = |i_1 i_2 \cdots\cdots i_K^* \cdots i_1^* \cdots i_N| \quad \text{双激发态} \qquad (2.170)$$

4. 使用“组态相互作用矩阵”表示分子 Hamiltonian

使用“组态相互作用矩阵”，最低本征值和相应的本征向量是借 Rayleigh-Schrödinger 微扰级数而计算的((2.157)式和(2.158)式)，激发态也可按类似方法处理。不过在具体实现以上两式的计算中，由于涉及简并而带来计算的困难，目前的 PCILO 计算多采用 Epstein-Nesbet 程序而将总的分子 $\mathscr{H}$ 分离为二。这样相当于取相互作用组态矩阵的对角部分作为未微扰 Hamiltonian 量 H，而非对角部分作为微扰 Hamiltonian 量 V 的表示。

2.2.6.3 PCILO 法的适用范围以及应用实例

PCILO 方法的实现，在目前基本上是依据QCPE(量子化学程序交换中心)提供的程序[59]，在能量上可达三级。第二级修正仅在负值上的加和((2.160)式)，并且值是很大的。因而它降到由充分 CI 应得到的相关能以下；第三级修正明显地小，这导致总的能量修正贡献为正，从而使能量比较接近真实值。它适用于 H 原子至 Cl 原子的含 s、p 轨道的除稀有气体以外的所有原子。但要求每个原子不得超过四配位，在组建杂化轨道时，不考虑 3 d 轨道，並要对所计算的分子必须写出 Lewis 结构式，以便构筑零级波函数。

该 PCILO 程序计算的价电子数可达 250 个，除了这个常用的在 CNDO/2 近似基础上的 PCILO 程序以外，还可获得 INDO 近似的 PCILO 程序 PCILINDO[60]，还有修改的 PCILO-GSPCILO 及包含过渡原子的 PCILO 2[61,62]。

PCILO 的计算通常以构象能图的形式给出，如能附带作图程序，这种构象能图的获得是很容易的。构象能图是两个连续的扭角(或称两面角)的函数，这应是经选择为主要影响构象的两个扭角，而其他扭角均按有机化学或生物化学的惯例加以固定或者直接引自 X-衍射结晶数据。

Pullman 等曾应用 PCILO 法计算二核苷酸的构象[63-65]，具体计算采用二糖单磷酸的模型(图 2-14)。

在实际计算中，旋转两个 P—O 键(P—$O_{3'}$和 $O_{5'}$—P)。这主要考虑到它们是主链骨架的主要扭曲轴。图 2-15 是二糖单磷酸两个扭角的构象能图，也是 PCILO法的典型的构象能图。

从图 2-15 可以看出，在 $\Phi_{P-O_{3'}} = \Phi_{O_{5'}-P} \approx 300°$ 附近，有一个总的极小值 ≈ 0 。这个

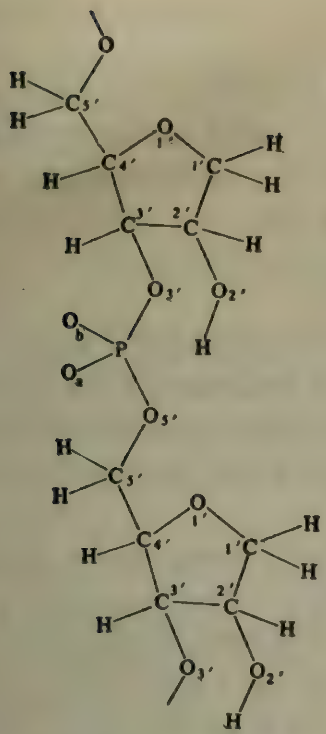


图 2-14 二糖单磷酸模型分子图

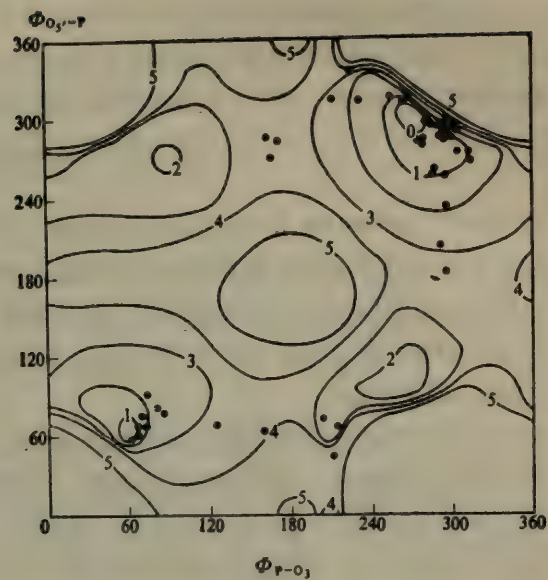


图 2-15 二糖单磷酸 PCILO 计算构象能图 kcal/mol*

值是该点的计算的构象能与最小构象能之差(kcal/mol),在构象能图上以等高线形式画出。300°附近的极小点常以 g^-g^-标出，且构象的稳定性面积很大，它们圈于 1 ～ 2 kcal/mol 的等势曲线之内。一个局部极小值位于 $\Phi_{P-O_3'}=\Phi_{O_5'-P}$ 在 60°～90°之间，记为 g^+g^+。此外，在 $\Phi_{P-O_3'}=\Phi_{O_5'-P}=180°$ 表示了一块高能平原（标以七七)，大约高出总极小值 5 kcal/mol。图 2-15 上的黑点是实验点，它们由二核苷磷酸和三核苷磷酸以及二甲基磷酸的 X 射线结晶学构象分析得到的。从图上看到 PCILO 计算预测值和实验观察的数据符合得很好。大多数实验测得构象落在稳定区 g^-g^- 和 g^+g^+ 之内或交界处，而无一落在七七高能平原区。

计算二甲基磷酸模型的 ab initio 法所获的结果完全类似于 PCILO 法[66]，两个具有相同稳定性极小值分别位于($\Phi_{P-O_3'}$，$\Phi_{O_5'-P}$)=(285°，285°）和(75°，75°）附近，並且在(180°，180°)附近区域高出总极小值约 7 kcal/mol。类似的结果也为 CNDO/2 法获得[67]。但若采用 EHT法[68],可预测出 7 个等价的能量极小区域的存在,但对这些极小的选择产生了较大的困难。这表明 EHT 法还不如 PCILO 法和其他较精确的量子化学方法的计算严谨可靠。当然这不是绝对的，也有不少成功的算例。

正因为 PCILO 法是分子轨道法范畴中计算速度快(按一般估计计算速度比同类分子 SCF

* 按国家标准局 1986 年发布的《中华人民共和国国家标准》，cal(卡）属于应废除的单位，在使用中应按 1 cal＝4.18 J(焦耳)换算成 J。

计算快 100 倍)、结果较为可靠的方法，因此，迄今仍是量子生物学中研究 生物分子和药物分子构象的主要量子化学方法[58]。

我国运用 PCILO 法研究构象问题起步较晚，1985 年才见报道[69]，且仅限于与 X-衍射实验相对比的工作，未涉及到生物大分子和药物分子。直至第三次全国量子化学会议上才有 PCILO 运用于研究生物分子的构象与活性关系的研究工作问世[70,71]。但是，可以相信，随着量子生物学的普及和深入，PCILO 法必将获得进一步的发展和应用。

2.2.7 分子力学计算法

分子力学(molecular mechanics)计算目前已发展成为结构化学的理论研究方 法之一[72]。这种基于分子经典力学的计算方法已从早期的有机化合物的分子结构和某些热力学参数为主，深入到研究有机化学反应机理、过渡态、溶剂化作用，结构与反应性，有机高分子结构，NMR 中的取代基效应等。在生物学领域中，应用该方法研究生物大分子的构象，药物结构与生物活性的关系，已呈方兴未艾之势。

分子力学计算是将能量直观地按物理学中一般力学的原理考虑了伸缩、弯曲、非键等能量项。由于应用这些概念可对分子的空间位阻、构象变化等问题进行定量的描述，而不涉及量子化学中的形成 Fock 矩阵和自洽叠代计算，因此在生物化学、药物化学领域受到广泛重视。

分子力学的另一个显著特点是计算速度快。我们知道，生物大分子和药物 分子的 结构活性关系中的一个重要的研究方面是分子的构象。用量子化学分子轨道法计算分子构象，可给出某些分子性质的较准确的结果。但这种自洽场的计算需化费大量的时间(计算机时按 原子 轨道数 N 的 3—4 次方的比例增加)，这就给实际计算带来了很大的障碍。但是，若采用分子力学计算方法，则化费的计算机时就会少得多(它仅与原子数目的平方成正比)，这就使得分子力学的计算在生物学领域内研究大分子构象中一显身手。

分子力学计算还可与经典的量子化学方法相辅相成，一方面可以用分子力学计算已优化好的构象所指示的分子空间坐标来进行量子化学计算，研究分子的电子结构信息。另一方面分子力学的一些经验参数除某些可从实验数值获得外，还需用量子化学从头计算法来优化参数。

正因为上述原因，分子力学计算在量子生物学领域中愈来愈受到人们的重视，其计算程序可在量子化学程序交换中心(QCPE)索取。

2.2.7.1 分子力学原理

分子力学原理首先在于考虑分子的经典力学的模型，构筑力场(force field)，把分子处理为由一套经典力学势能函数支配的原子的排列。这个力场的势能应包括非键和价键的相互作用，通过不同势能函数的解析表达式，它们之间的总和即为分子的构象能。这个力场的势能 函 数一般包括键的伸缩、键角的弯曲、绕价键旋转的内旋转、Van der Waals 作用力及静电和偶极的相互作用。其势能可写成

$$V=\sum V_{stretch}+\sum V_{bend}+\sum V_{torsion}+\sum V_{dipol}+\sum V_{charge}+\sum V_{VDM} \tag{2.171}$$

式中 $V_{stretch}$ 是键伸缩势能；V_{bend} 是角弯曲能；$V_{torsion}$ 系内旋转能；V_{dipol} 是偶 极 相 互 作 用 能；

V_{charge} 是静电相互作用能；V_{VDM} 是 Van der Waals 作用力；每项分别对所有原子求和。

基于上述经典力学模型，已有不少采用不同的势能函数表达式和相应参数的不同的力场方案问世，各有侧重，各有特长，并各自有其相应的计算机程序。其中应用最广影响最大的首推 Allinger 等建立并系列化的 MM 1、MMP 1、MM 2、MMP 2 等方法中推演的势能解析式和相应的参数，其计算机程序 MM 2 和 MMP 2 也以不同机型的版本在各国广泛地应用。

以下以 Allinger 建立的 MM 2 方法中的势函数为主分别就(2.171)式中的各项具体解析式作介绍[78]，由于势能的线性加和性，以及为简要地叙述力场的构成，有必要首先讨论两个原子间的势能。

1. 键伸缩和角弯曲势能

键伸缩和角弯曲可由一般的 Morse 曲线来描述，我们可以设想原子之间用弹簧联结起来，构成了分子，在近平衡态处，可用 Hook 定律来表示键长键角偏离其“标准”时所增加的能量。

$$V_{str.}=\frac{1}{2}K_r(r-r_0)^2 \tag{2.172}$$

$$V_{bend}=\frac{1}{2}K_\theta(\theta-\theta_0)^2 \tag{2.173}$$

式中 K_r 是键伸缩力常数，K_θ 是键角弯曲力常数，它们通常由拟合小分子振动光谱的实验数据而获得。r 是键长，r_0为这个力场模型下的标准键长或称“本征”键长，θ 是键角，θ_0 是标准键角。上式应对分子内所有键长、键角求和。

为了更精确地计算，时常对 $V_{str.}$ 和 V_{bend} 进行校正，添加高次项。表 2-8 和表 2-9 列出了二个力场的参数。

表 2-8 $V_{str.}$ 计算参数

力场模型作者	C—C		C—H		
	K_r	r_0	K_r	r_0	K'_r
Allinger	4.40	1.523	4.60	1.113	2.0
Bartell	2.34	1.166[a]	3.85	1.0203[a]	2.0
		1.534[b]		1.1068[b]	

a 仅在二次项时使用，b 仅在三次项中使用

表 2-9 V_{bend} 计算参数

力场模型作者	C—C—C		C—C—H		H—C—H			
	K_θ	θ_0	K_θ	θ_0	K_θ	θ_0	K'	n
Allinger	0.450	109.5	0.360	109.4	0.320	109.4	7.10^{-8}	4
				110.0				
Bartell	0.629	109.47	0.322	109.47	0.350	109.47	0.0	—

$$V_s=\frac{1}{2}K_r[(r-r_0)^2-K_r'(r-r_0)^3] \tag{2.174}$$

$$V_b=\frac{1}{2}\cdot 3.046\cdot 10^{-4}\cdot K_\theta(\theta-\theta_0)^2+K'_\theta(\theta-\theta_0)^n \tag{2.175}$$

为进一步改进分子力学模型，往往要引进伸缩-弯曲交叉项。这是因为当键伸缩时，常常伴随着键角的减小，若分子有较大变形性，如环烷，这种互相影响就较显著。这种影响主要存在于1,3原子间的非键作用，普通力场并不包含这种作用，故而必须考虑伸缩-弯曲交叉项，可由下式求得

$$V_{sb}=\sum_{i<j}\frac{K_{ij}^{sb}}{2}(r_i-r_i^0+r_j-r_j^0)(\theta_{ij}-\theta_{ij}^0) \tag{2.176}$$

式中 K^{sb} 为伸缩-弯曲参数，θ 和 θ^0 为实际键角和平衡键角，r_i 和 r_j 为键 i 和键 j 的键长，r_i^0 和 r_j^0 是键 i 和键 j 的键长参数。

2. 扭转作用势能

扭转作用势能就是围绕键转动而产生分子能量的变化。绕键旋转在分子力场中是以“扭角”ω 表示 (torsional angle)，事实上它与量子化学计算中确定分子空间坐标的二面角 (dihedral angle)同义。

扭转作用势能的计算，常以 Fourier 级数展开式表示

$$V_{tor.}=\frac{1}{2}V_1(1+\cos\omega)+\frac{1}{2}V_2(1-\cos 2\omega)$$

$$+\frac{1}{2}V_3(1+\cos 3\omega)+\cdots\cdots \tag{2.177}$$

对许多有机和生物分子而言，前三项展开已足够精确，上式中 V_1、V_2、V_3 是参量，它是由事先选择一个分子使其预测的构象符合实验结果而拟合的常数。

当分子的构象发生变化时，往往扭角变化引起势能 V 改变要比由伸缩和角弯曲的改变大得多。在处理生物分子等较大分子时，时常保持键长、键角不变，仅改变扭角。

3. 静电作用势能

含有极性基团的分子中基团的电荷相互影响产生静电作用势能。可由下式表示

$$V_{charge}=\frac{q_i\cdot q_j}{D\cdot r_{ij}} \tag{2.178}$$

式中 q_i 和 q_j 分别为原子或极性基团 i 和 j 的电荷，D 是有效介电常数。

有的力场模型还考虑了分子内氢键形成而引起静电作用势能的变化[74]，无疑这对生物分子中诸如氨基酸、蛋白质的分子力学计算是很必要的。

针对生物分子的力场模型，对于多肽、蛋白质中—CO—及—NH_2—之间的分子内氢键用下式计算 V_H

$$V_H=4e\left[\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^6\right]-\frac{\mu_{CO}\cdot\mu_{NH}}{r^3}$$

$$\cdot(2\cos\theta_0\cdot\cos\theta_H-\sin\theta_0\cdot\sin\theta_H\cdot\cos\alpha) \tag{2.179}$$

式中 θ_0 是 C=O 键的偶极矩 μ_{CO} 与氢键方向的夹角，(图 2-6)，θ_H 是 N—H 键偶极矩 μ_{NH} 与氢键方向的夹角，α 是包含 C、O、H 的平面与包含 N、H、O 的平面的二面角。

4．偶极相互作用势能

偶极相互作用考虑到静电-偶极相互作用和偶极-偶极相互作用。其值可由下式获得

$$V_{dipol}=\frac{\mu_i\cdot\mu_j}{D\cdot r_{ij}^3}(\cos\chi-3\cos\alpha_i\cdot\cos\alpha_j) \tag{2.180}$$

式中 μ_i、μ_j 为偶极矩，χ 为两偶极矩间的夹角。α_i、α_j 分别为偶极矩与键中点联线间的夹角，r_{ij} 是键中点间距(图 2-17)。

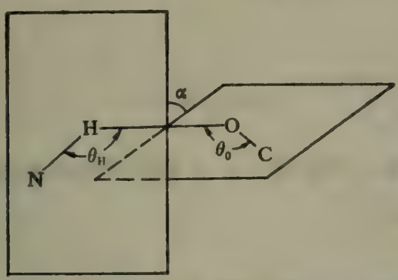


图 2-16

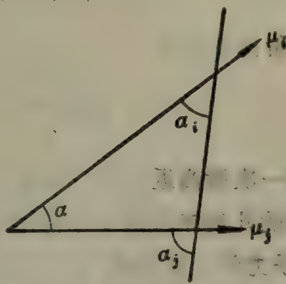


图 2-17

5．Van der Waals 作用势能

考虑到非键原子间的作用，势能中还必须包含 Van der Waals 作用，在确定参数时，假定分子内与分子间非键原子间作用遵循同样的规律。

有两种势函数可以描述 Van der Waals 作用势

Lennard-Jones 函数

$$V_{VDM}=\varepsilon\left[\left(\frac{r_0}{r}\right)^{12}-2\left(\frac{r_0}{r}\right)^6\right] \tag{2.181}$$

Buckingham 函数

$$V_{VDM}=\frac{\varepsilon}{1-\frac{6}{\alpha}}\left[\frac{6}{\alpha}\ e^{\alpha/(1-r/r_0)}-\left(\frac{r_0}{r}\right)^6\right] \tag{2.182}$$

这两种函数吸引部分都是 r^{-6}，但对排斥部分的处理不一样。L-J函数采用 r^{-12}，而 Buckingham 以指数形式的斜率增加方式表示。

目前最常用的还是 Allinger 基于 Hill 方程采用的 Buckingham 函数解析式

$r_0/r<3.311$

$$V_{VDM}=K_r\left[2.90\times10^5\cdot e^{-12.50r_0/r}-2.25\left(\frac{r_0}{r}\right)^6\right] \tag{2.183}$$

$r_0/r\geqslant3.311$

$$V_{VDM}=336.176\left[K_r\cdot\left(\frac{r_0}{r}\right)^2\right] \tag{2.184}$$

式中 r_0、r 分别为 Van der Waals 半径与原子间距，K_V 为两个原子硬度参数的几何平均

$$K_r=(E_i-E_j)^{1/2} \tag{2.185}$$

上述数项势能之和在分子力场中称为立体能(steric energy)。求分子稳定构象或稳定态就需要采用优化的方法，使其立体能最小，所以这个优化过程在数学中属多因素最优化方法。Allinger 采取其中的 Newton-Raphson 优化法，亦称块对角矩阵NR 优化法。

优化能量一般公式为

$$\frac{\partial V}{\partial x_i}=0 \tag{2.186}$$

在 NR 优化法中，计算势能对几何参数的一次微商和二次微商（力常数），使坐标 $x^{(K+1)}$ 比 $x^{(K)}$ 更接近于平衡的几何参数，逐次计算

$$x^{(K+1)}=x^{(K)}-H^{-1}(x^{(K)})\nabla V(K) \tag{2.187}$$

式中 $H(x^{(K)})$ 是二阶微商矩阵

$$H(x^{(K)})=\left(\frac{\partial^2 V(K)}{\partial x_i \partial x_j}\right) \qquad (1\leqslant i,j\leqslant n) \tag{2.188}$$

采用叠代法使所有一次微商近为零（小于一个相对极小限定值），计算所得的构象即为最后平衡构象。其所指示的几何参数即为平衡时的空间坐标。

2.2.7.2 分子力学计算流程

一般的分子力学计算的实现，可以用图 2-18 所表示的流程框图作个梗概的了解。

2.2.7.3 分子力学在量子生物学中的应用

分子力学程序大部分可由 QCPE 获得[72]。目前用得最广的是 Allinger 发展的 MM 2 和 MMP 2 程序。它们分别由 MM 1 和 MMP 1 改进而来。MM 2 版本可用于计算烷烃、非共轭的烯烃、炔烃、非共轭的羰基化合物、醇类、醚类、硫化物、硫醇、单取代卤代物和胺类。MMP 2 还可用于含有离域的 π 电子体系，这对于药物的计算非常适用。

Allinger 在 1982 年写的第一本关于分子力学的专著[73]，较全面地总结了 1980 年以前的有关生物大分子的构象和药物的化学结构与生物活性之间的相互关系的研究[74]。如生物分子ATP构象的分子力学计算[75]，对以后的大分子计算产生了积极的影响。

近年来，分子力学的研究获较快发展，继Allinger的专著以后，又陆续有不少评述问世[76-79]。在生物分子和药物分子的分子力学计算中有其广泛的用途[80-89]，特别是 QCPE，可获得 MM 2 和 MMP 2 分子力学的程序以后，在药物和生物分子的应用方面更为令人鼓舞[90-92]。

除了 Allinger 等发展的 MM系列的分子力学计算程序以外，各国学者还结合本身的研究对象，各自发展了不同模型的分子力学计算程序。Kollman 等建立了 AMBER 程序系统[86,93-97]，Rao 和 Kollman 用这个系统进行了一系列生物分子的构象研究，这个程序系统在生物分子的应用中相当成功，故而影响甚广。药物分子的分子力学计算中，Andrews等建立了称之为经典势能计算(classical potential energy calculation)[98,99]的分子力学方法，并结合绘图程序建立了 COMOL 程序系统，对不同作用机制的药物构象与受体的结合从理论上作了较深入的探讨[100,101]，已经引起量子药理学研究者的高度关注。值得注意的是 Pullman 等建立的分子力学计算法 SIBFA 和 SIBFA 2，它们将分子力学与量子化学紧密地有机地结合起来。这些分子力学方法称为从头算碎片间相互作用加和法(sum of interactions between fragments computer by ab

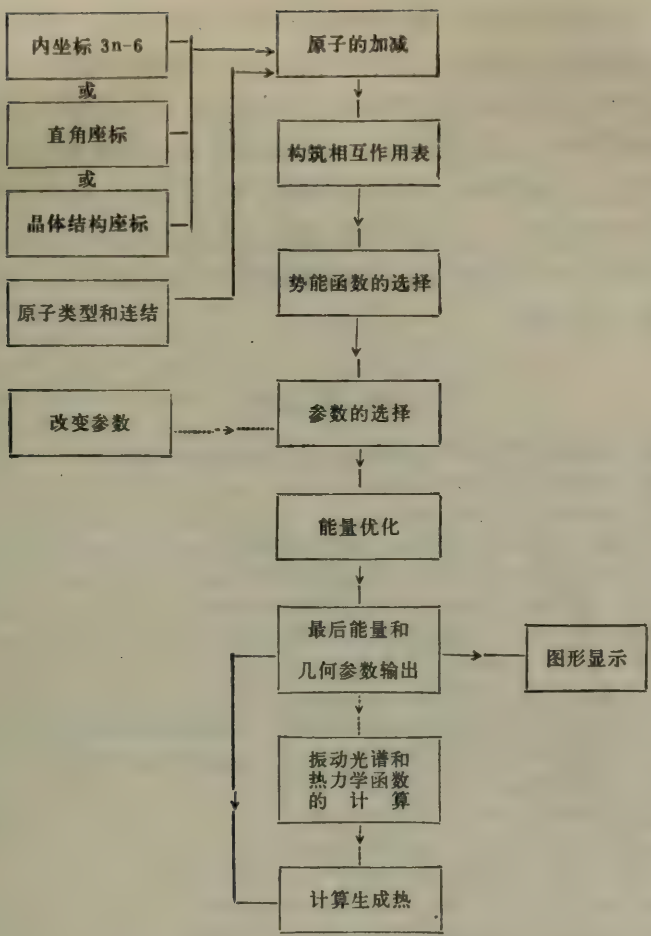


图 2-18 分子力学计算流程框图

initio，简称 SIBFA)[102]。这个方法以分子碎片的从头计算为基础，得到各原子的一系列特征数据；进而依据分子力学的一系列公式，计算由于分子构象变化而引起的分子内相互作用能。陈凯先等用这个方法计算 Daunomycin 及其衍生物与核酸的作用，研究这些抗癌药物与 DNA 的作用形式。已发表的工作表明，SIBFA 对于研究嵌入型抗肿瘤药物与 DNA 的作用是一个有用而方便的工具[103-107]。

国内的分子力学研究工作开展不算太晚。江寿平和杨频等[107]在量子生物学领域中，应用分子力学探讨生物大分子的构象，在国内处于领先水平。李树森等在研究有机化学反应和研究有机磷化合物的结构与萃取稀土元素的性能方面所作的工作已属开创性的探索[108]。随着分子力学程序的普及和推广，这方面的研究必将展示广阔的前景。

以上对迄今为止常用于量子生物学领域的主要计算方法的原理作了概括的介绍。我们认为，任何一种方法都不是完美的，它都或多或少地有不可避免的缺陷。如 HMO 法简单易行，却又嫌

精确度不够;ab initio 虽然近似最少,但耗费机时很多,占用内存很大,要视计算机的功能、支付费用的能力和研究对象是否必需而定。分子力学法进行构象研究速度快,但必竟基于经典力学,时常用 PCILO 或 CNDO 法优化构象互相配合。所以任何一个方法都不应偏废。应当分析计算对象,区别对待。有时用 HMO 法可获得较好的结果,就不必一定用更高层次的计算方法(至少就目前国内的情况而言)。应该看到,量子生物学的发展,反过来又促进各种方法的改进和发展。目前,可以方便地计算生物大分子(如蛋白质)的量子化学计算方法尚无重大突破。然而,可以预料,在量子化学计算方法学研究者和量子生物学研究者的通力协作下,量子化学计算在生物学中的应用必将获得进一步发展。

参考文献

[1] W. G. 理查兹著,卓济苍等译:《量子药理学》,科学出版社,(1981)。

[2] 谢雷鸣、吴吉安:化学通报,(4):14(1981).

[3] 刘次全等:分子科学学报,1(2):19(1981).

[4] Coulson, C. A. et al.: Proc. Roy. Soc., A 191,34(1947).

[5] Coulson, C. A.: Discussion Faraday Soc., 2:9(1947).

[6] Jordan, F. G. et al.: Theor. Chim. Acta, 15:211(1969).

[7] 朱永等:《量子有机化学》(上),上海科学技术出版社,(1983).

[8] 戴乾圜:中国科学,(10):964(1979).

[9] Hoffman, R.: J. Chem. Phys., 39:1397(1963).

[10] J. A. 波普尔著,江元生译:《分子轨道近似方法理论》,科学出版社,(1976).

[11] 王志忠等:分子科学与化学研究,2:47(1982).

[12] 邵俊、吴吉安等:上海科学技术大学学报,2:45(1985).

[13] 宣建成、江寿平:中国科学(B辑),(7):681(1986).

[14] Loew, G. H. et al.: Mol. Pharm., 26:19(1984).

[15] Loew, G. H. et al.: Cancer Biochem. Biophys., 7:109(1984).

[16] Koryckwa, K. et al.: J. Am. Chem. Soc., 107:4273(1985).

[17] 廖沐真等:《量子化学从头计算法》,清华大学出版社,(1984).

[18] Hellman, H.: J. Chem. Phys., 4:324(1963).

[19] Ohata, O. et al.: J. Phys. Soc. Japan, 21:2306(1966).

[20] Hehre, W. J. et al.: J. Chem. Phys., 51:2657(1969).

[21] Chen Kaixian et al.: Kexue Tongbao (China), 31:1464(1986).

[22] 陈凯先、嵇汝运:分子科学学报,1:103(1982).

[23] Bonifacic, V. et al.: J. Chem. Phys., 62:1509(1975).

[24] 陈凯先、嵇汝运等:分子科学学报,1:65(1981).

[25] Lin, J. H. et al.: J. Phys. Chem., 84:642(1980).

[26] Gresh, N. et al.: Theor. Chim. Acta, 66:1(1984).

[27] Gresh, N. et al.: Inter. J. Quant. Chem., 29:101(1986).

[28] Chen Kaixian et al.: Nucleic Acids Res., 14:3799(1986).

[29] Slater, J. C.: Phys. Rev., 81:385(1951).

[30] Segal, A.: "Semiempirical Methods of Electronic Structure Calculation", (Plenum Press, New York, london), (1977).

[31] Johnson, K. H.: J. Chem. Phys., 45:3085(1966).

[32] 唐敖庆等著:《量子化学》,科学出版社,(1982).

[33] 李俊清:《量子化学中的 X_α 方法及其应用》,安徽科学技术出版社,(1984).

[34] 唐敖庆:有机化学,(2):156(1985).

[35] Politzer, P.: "Chemical Applications of Atomic and Molecular Electrostatic Potentials",

Politzer, P. and Truhlar, D. G. (Ed), (Plemum Press, New York and London), (1981).
[36] Kaufman, J. J. et al.: ibid 335.
[37] Pullman, A. et al.: ibid 381.
[38] 陈凯先、嵇汝运等：分子科学与化学研究，3(1):25(1983).
[39] 陈凯先、吴吉安、嵇汝运：分子科学与化学研究，3(3):1(1983).
[40] 吴吉安、陈凯先、嵇汝运：华西医药杂志，(1):71(1986).
[41] Pretter, L. G. et al.: Theor. Chim. Acta, 25:83(1972).
[42] 吴吉安、陈凯先、嵇汝运等：«全国第二届量子化学年会论文集»，(吉林长春)，(1983).
[43] Xu Shujuan et al.: J. Mol. Sci., (Wuhan, China), 4:11(1986).
[44] 郭直惟等：华西医药杂志，1:74(1986).
[45] Bohl, M. et al.: Eur. J. Med. Chem., 21:193(1986).
[46] Buydens, L. et al.: Eur. J. Med. Chem., 21:35(1986).
[47] 陈凯先、吴吉安、嵇汝运：中国科学(B辑)，(12):1257(1986).
[48] Giessner-Prettre, C. et al.: Theor. Chim. Acta, 33:91(1974).
[49] Andrews, P. R. et al.: Inter. J. Quant. Chem. QBS., 7:293(1980).
[50] Petrongolo. C. et al.: J. Med. Chem., 20:1645(1977).
[51] Macchia, B. et al.: Eur. J. Med. Chem., 15:515(1980).
[52] Van der Waterbeemd, H. et al.: J. Med. Chem., 29:600(1986).
[53] Diner, S. et al.: Theor. Chim. Acta, 13:1(1969).
[54] Diner, S. et al.: ibid, 15:100(1969).
[55] Malrieu, J. P. et al.: ibid, 13:18(1969).
[56] Malrieu, J. P. et al.: ibid, 15:110(1969).
[57] Pullman, B.: "Quantum Mechanics of Molecular Conformations", Pullman, B. (Ed), (John Wiley & Sons Ltd London), (1976).
[58] Richards, W. G.: "Quantum Pharmacology", Second Edition, (Butterworth &Co Ltd London), (1983).
[59] QCPE, 220,221.
[60] QCPE, 371.
[61] QCPE, 272.
[62] QCPE, 390.
[63] Pullman, B. et al.: Biochim. Biophys. Acta, 269:1(1972).
[64] Perahia, D. et al.: "Conformation of Biological Molecules and Polymers", 5 th Jerusalem Symposium on Quantum Chemistry and Biochemistry, (Bergmann, E. D. and Pullman, B.Eds.), (1973).
[65] Perihia, D. et al.: Biochim. Biophys. Acta, 340:299(1974).
[66] Newton, M. D.: J. Am. Chem. Soc., 95:256(1973).
[67] Tewari, R. et al.: J. Theor. Biol., 46,229(1974).
[68] Saran, A. et al.: J. Theor. Biol., 33:407(1971).
[69] 白春礼等：中国科学(B辑)，4:289(1985).
[70] 徐筱杰等：«全国第三届量子化学会议论文集»，(成都)，(1987).
[71] 关玥等：«全国第三届量子化学会议论文集»，(成都)，(1987).
[72] Burkert, U. et al.: "Molecular Mechanics", ACS Monograph 177, American Chemical Society, Washington, D. C. (1982).
[73] Allinger, N. L. et al.: J. Am. Chem. Soc., 93:1637 (1971); Tetrahedron, 30: 1579 (1974); J. Am. Chem. Soc., 95:3893 (1973); ibid, 96:5100 (1976); Adv. in Phys. Org. Chem., 13:1 (1976).
[74] 刘若庄等：分子科学与化学研究，2:121(1982).
[75] Millner, O. E. et al.: Biopolymers, 14:2159(1975).
[76] Osawa, E.: Top. Stereochem., 13:117(1982).

[77] Simonertta, M.: Inter. Rev. Phys. Chem., 4:39(1985).
[78] Favini, G.: Theochem., 10:139(1983).
[79] Meyer, A. Y.: Chem. Halides, Pseudo-Halides Azidics, 1:1(1983).
[80] Squier, G. J.: Eur. J. Med. Chem., 21:213(1986).
[81] Froimowitz, M. et al.: J. Comput. Chem., 5:291(1984).
[82] Kollman, P.: P. Acd. Chem. Res., 18:105(1985).
[83] Hopfinger, A. J.: J. Med. Chem., 26:990(1983).
[84] Cory, E. J. et al.: Tetrahedron Leff, 25:4325(1984).
[85] Remers, W. A. et al.: J. Med. Chem., 29:1256(1986).
[86] Kollman, P. A.: J. Am. Chem. Soc., 106:1765(1984).
[87] Caldwer, J. et al.: Biopolymers, 25:249(1986).
[88] Hall, D. et al.: J. Compt. Chem., 5:441(1984).
[89] Aqarwal, A. K. et al.: J. Biomol. Struct. Dyn., 1:873(1983).
[90] Froimowitz, M. et al.: J. Med. Chem., 29:573(1986).
[91] Griffin, J. F. et al.: J. Med. Chem., 29:778(1986).
[92] Froimowitz, M. et al.: ibid, 29:1570(1986).
[93] Pattabiraman, N. et al.: Biopolymers, 26:403(1987).
[94] Weiner, S. J. et al.: J. Comput. Chem., 2:287 (1981).
[95] Weiner, S. J. et al.: J. Am. Chem. Soc., 106:765(1984).
[96] Weiner, S. J. et al.: "AMBER (UCSF) Version 2.0 Dept. of Pharmaceutical Chemistry School of Pharmacy University of California", (San Francisco California 94143 U. S. A.), (1985).
[97] Rao, S. N. et al.: Biopolymers, 25:267(1986).
[98] Giglio, E.: Nature, 222:339(1969).
[99] Andrews, P. R. et al.: J. Med. Chem., 28:293(1985).
[100] Lloyd, E. J. et al.: ibid, 29:453(1986).
[101] Wong, M. G. et al.: ibid, 29:562(1986).
[102] Gresh, N. et al.: Theor. Chim. Acta, 66:1(1984).
[103] Chen Kaixian et al.: Mol. Pharmacol., 30:279(1986).
[104] Chen Kaixian et al.: J. Biomol. Struct. Dynam., 3:445(1985).
[105] Chen Kaixian et al.: Nucleic Acids Res., 14:2251(1986).
[106] Chen Kaixian et al.: Biopolymers, 26:87(1987).
[107] 郝鸣鸿等：化学通报,(11):14(1984).
[108] Shusen, L. et al.: J. Mol. Sci. (Wuhan, China), 5:73(1987).

刘次全
吴吉安
曹　槐
王　莹

第3章 原子间相互作用和分子间互相作用[1-2]

在上一章讨论量子生物学的理论方法和理论指数的基础上，在这一章里，我们将着重介绍有关化学键，Van der Waals 相互作用，氢键相互作用，生物分子的特异作用和识别等问题。

我们知道，生物分子以及由它们所组成的体系，其结构与功能是密切相关的。然而，维持生物分子特定的结构又取决于分子内和分子间的各种相互作用。不仅如此，生物分子内原子间相互作用和分子间的相互作用(相互作用力)又是了解特异作用和识别机理的根据。因此，研究生物分子内原子间和分子间的各种作用力的性质、特征，及其与结构的关系，尤其是某些作用力的半定量和定量处理，自然就成了量子生物学的基本内容。回顾近几年来量子生物学的进展，对说明生物分子中原子间相互作用和分子间的相互作用以及生物分子的化学性质，确实起了较大的作用。但是，这些研究工作似乎还缺乏对生物分子这一特殊对象的整体认识，如能对此进行更透彻的探讨，就可以取得更大的进展。

3.1 生物分子中原子间的相互作用

关于生物分子中原子间相互作用的问题，我们认为归根结底主要是讨论化学键的问题(当然也还包括其它的一些问题)。那么，量子力学对于理解化学键有什么用处？对这样的问题，确实是难以回答的。但是，从另一方面讲，首先真正地、演绎地、严格地说明化学键的，主要还应归功于量子力学。当然，即使是在今天，也还不能简单地认为对化学键的理解已经很充分了。因为，事实上，对于“化学键的本质”这个问题(参见文献[3])，即使把有关的物理学家和化学家集中起来讨论上几天也未必能获得什么明确的结果。相反地，倒是会遗留下很多无法了结的问题。何况是对于具有种种生物活性的生物大分子中原子间相互作用的讨论。

3.1.1 相互作用力和能量

分子中的原子之间，或者分子与分子之间的作用力有许多不同的名称，但基本上都属于同一类型，即带电粒子之间的静电引力，这种力在生物分子以及大而复杂的体系(蛋白质和核酸等生物分子本身就是复杂的多聚高分子体系)中起着主要作用。

大量的研究表明，要直接测定两个微观对象(原子、分子)之间的相互作用力是不可能的。但在许多情况下常常可以测出相互作用的能量。例如，当 Na^+Cl^- 在 Na^+ 与 Cl^- 相距无限远时，可以认为二者之间没有相互作用，因而可以规定此体系的势能为零。但是，当这两个粒子由于引力作用而彼此接近到相距为 r 时，其势能应为

$$U_1=\int_{\infty}^{r}\frac{q^2}{r^2}\mathrm{d}r=-\frac{q^2}{r} \tag{3.1}$$

式中，U_1 表示电量各为 $\pm q$ 的离子相距为 r 时的势能值。从(3.1)式中可以看出，U_1 是负值、且 r 越小，U_1 值越小。然而，当两个离子相距很近时，原子核之间的斥力开始起作用，距离越近，斥力越大。由于原子核之间的斥力而产生的势能是按指数规律变化的，可表示为

$$U_2 = be^{-r/a}$$

式中 b 和 a 均为常数。整个体系的势能 U 可表示为 U_1 和 U_2 之和，即

$$U = U_1 + U_2 = -\frac{q^2}{r} + be^{-r/a} \tag{3.2}$$

U 与 r 的关系称为势能曲线。图 3-1 为离子化合物 NaCl 离子键的势能曲线。由图可以看出，当距离很近时，斥力势能占优势，U 值为正，而 U 由负过渡到正有一极小值，这是在引力与斥力相等时，体系处于最稳定状态。对于 NaCl，此最低势能值约为 -120 kcal/mol。

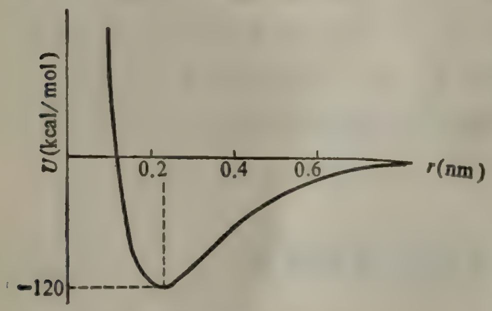


图 3-1 NaCl 离子键的势能曲线

所谓键能（解离能），即是使两个原子核从最稳定位置（能量最低的 r）分离到无限远处所需的能量。由于键能可以通过各种物理的和化学的方法测定，因而用相互作用能来表示原子间的相互作用比用相互作用力更有意义。

在实际处理问题时，为方便起见，通常把相互作用分成两类，即强相互作用——强力（属于强相互作用的有共价键和离子键等）和弱相互作用——弱力（属于弱相互作用的有 Van der Waals 力和氢键等）。前者通常维持着分子的基本结构；后者在量子生物学中是不可忽视的一类力，在维持生物大分子的构象和功能活性中起着重要作用。

强相互作用是使分子中的原子之间结合在一起的主要相互作用。这些作用决定着生物大分子的一级结构。由表 3-1 可以看出，强相互作用的结合能（键能）大大超过了分子的平均热运动能。我们知道，作用力的强弱在这里是相对于平均热能 kT 而言的，其中 k 是通常的 Boltzmann 常数，T 是绝对温度。在正常体温下，该平均热能的数量级是 2.5×10^{-2}eV/粒子，相当于 0.6kcal/mol。

表 3-1 一些典型键的解离能

键	解离能	
	eV	kcal/mol
H—H	4.45	100
C—C	2.55	58
C—H	3.80	86
C—N	2.13	48
C—N(肽)	3.03	70
C=C	4.35	99
C=O	6.30	143
C≡C	5.35	122

弱相互作用在数值上虽然比强相互作用小得多，但在维持生物大分子的二级、三级、四级结构中却起着相当重要的作用。正因为这类相互作用较弱，因此破坏这些相互作用所耗费的能量也较小。当生物分子受热时，这些弱相互作用将首先受到破坏，并以四级、三级和二级结构这样的顺序丧失生物大分子的高级结构，从而对生理功能造成明显的影响。

3.1.2 原子间相互作用的分类

正如 3.1.1 中所述，原子间的相互作用力基本上都属于带电粒子间的静电作用力。因此，我们需要区别以下两种情况：

(1) 两个相互作用的粒子的荷电状态如何？是中性原子还是离子？

(2) 相互作用的原子或离子的电子结构如何？即需要区分它们的最外电子层是闭壳层（全满）还是开壳层（未满）。因为闭壳层原子之间的相互作用力和开壳层原子之间的力差别甚大。

研究表明，开壳层和闭壳层体系的行为是截然不同的，据此，我们就可以对原子间的相互作用力作如下的分类：

(1) 闭壳层阳离子和闭壳层阴离子之间的相互作用。如 Na^+和 Cl^-，Na^+和 O^{2-}等。

(2) 闭层壳的中性原子间的相互作用。如 He—He 等（属 Van der Waals 相互作用）。

(3) 开壳层离子（一般是阳离子）和一个以上的闭壳层离子（一般是阴离子）之间的相互作用。如生物金属离子和配位体阴离子所形成的配合物（参见第十二章）。

(4) 两个都是开壳层的中性原子之间的相互作用。例如 H—H（属共价键类型）等。

需要指出，这样的分类虽然比较方便，但并不严格，尤其在处理生物分子这样的复杂体系时更是如此。不过，尽管这样的分类比较粗略，但对我们探讨原子间相互作用力，仍是有益的。

3.1.3 定域键和离域键

正如前面所述，我们认为有关分子中原子间相互作用的问题主要是讨论化学键的问题。具体地说，在量子生物学中，我们关心的是多原子分子。然而，多原子分子中的每一个电子是在更多的原子核和其它电子形成的场中运动的，情况比较复杂。但是，我们可以把这些电子在分子中的成键情况分为两类：一是限于一对原子之间的所谓定域键（localized bond），大多数分子属于这种情况。二是电子可以在三个或更多的原子核之间运动的所谓离域键（nonlocalized bond）。实践说明，这是简化问题的一个好的近似方法。因为，即使在离域作用很突出的分子中（如苯等），定域键也是在分子形成中起骨架作用的。因此，下面我们先讨论定域键的问题。

3.1.3.1 定域键

我们知道，原子之间或分子之间发生相互作用时，分子内和分子间的电子分布将会发生改变。一些实验事实表明，分子中原子间相互作用的某些性质主要同与键相连的原子有关，这说明成键电子密度的分布，在一定程度上有定域分布的特点。每个键相当于在一个键轨道中的一对电子，这时电子主要定域在这键的两个原子之间。

然而，在实际处理生物分子中原子间的相互作用时，又不可避免地要涉及到构成生物分子的六种基本元素原子的电子组态问题〔这六种基本元素是：氢（^1H）、碳（^6C）、氮（^7N）、氧（^8O）、磷（^{15}P）

和硫(^{16}S)〕。因为电子组态问题在生物分子中是不可忽视的。

就氢原子而论,其基态电子组态为$(1s)^1$。

碳原子在基态时的电子组态为$(1s)^2(2s)^2(2p)^2$。其中,4个电子分别处于$1s$和$2s$,而且它们已经配对故不参与形成化学键。但是,有两个未充满的$2p$轨道各包含一个未成对电子。这似乎表明,当一个碳原子跟其它原子(包括其它碳原子)发生相互作用形成化学键时,它应该是二价的。但实验事实却说明碳原子是四价的,且四个化学键是等同的,它们的键角为109°28′。为了使理论与实验结果相一致,可以假定一个$2s$电子进入到空的$2p$轨道,从而给出$(1s)^2(2s)^1(2p)^3$组态。如果假定一个键是由$2s$电子形成,三个键是由$2p$电子形成,那么显然地与四个键是等同的实验结果相冲突。为此,Pauling提出了L壳层中的四个原子轨道线性组合形成sp^3杂化轨道,其形式为

$$\mathrm{b}_i(C_{2s})+\mathrm{d}_i(C_{2p_x})+\mathrm{e}_i(C_{2p_y})+\mathrm{f}_i(C_{2p_z})$$
$$i=1,\cdots,4 \tag{3.3}$$

为了获得最大的重叠,令每个函数的极大值处于四面体的一个顶点。若以甲烷为例,我们可以这样来解释,即如果把四个轨道按照下列方式进行线性组合(波函数ϕ_{2s}、ϕ_{2p_x}、ϕ_{2p_y}和ϕ_{2p_z}的线性组合在能量上是有利的),于是可以得到四个等同的轨道:

$$\begin{aligned}
\phi_1&=\frac{1}{2}(\phi_{2s}+\phi_{2p_x}+\phi_{2p_y}+\phi_{2p_z})\\
\phi_2&=\frac{1}{2}(\phi_{2s}-\phi_{2p_x}+\phi_{2p_y}-\phi_{2p_z})\\
\phi_3&=\frac{1}{2}(\phi_{2s}+\phi_{2p_x}-\phi_{2p_y}-\phi_{2p_z})\\
\phi_4&=\frac{1}{2}(\phi_{2s}-\phi_{2p_x}-\phi_{2p_y}+\phi_{2p_z})
\end{aligned} \tag{3.4}$$

这四个sp^3杂化轨道间的夹角等于109°28′,跟实验值完全相符。

在甲烷中每个键可以看作是碳原子上一个sp^3杂化轨道和一个氢原子的$1s$轨道重叠而形成的。很明显,形成四个键而不是二个键所得到的能量足以补偿把一个电子从$2s$轨道激发到$2p$轨道所需的能量。

为了弄清构成分子时原子间的结合是如何建立的,我们必须从键合来源于两个原子轨道的最大重叠这个原则出发。下面继续讨论乙烯和乙炔的情形。

实验表明,乙烯分子是平面型的,所有的键角接近于120°,而且它具有由分子式$CH_2{=}CH_2$中的双键所表示的不饱和性。如果一个$2s$轨道和两个$2p$轨道杂化形成三个sp^2杂化轨道,则这三个轨道可以写作:

$$\phi_1=\frac{1}{\sqrt{3}}\phi_{2s}+\frac{\sqrt{2}}{\sqrt{3}}\phi_{2p_x}$$

$$\phi_2=\frac{1}{\sqrt{3}}\phi_{2s}-\frac{1}{\sqrt{6}}\phi_{2p_x}+\frac{1}{\sqrt{2}}\phi_{2p_y}$$

$$\phi_3=\frac{1}{\sqrt{3}}\phi_{2s}-\frac{1}{\sqrt{6}}\phi_{2p_x}-\frac{1}{\sqrt{2}}\phi_{2p_y} \tag{3.5}$$

由(3.5)式可以推出，这三个轨道将位于同一平面上并互成 120°，即轨道对称轴指向等边三角形的三个顶点。余下的一个 2 p_z 轨道将与三个 sp^2 杂化轨道所组成的平面垂直。

在乙炔分子中，一个 s 轨道和一个 p 轨道杂化时，将形成两个互成 180 °的 sp 杂化轨道，其余二个未杂化的 2 p_y 和 2 p_z 轨道都和 sp 杂化轨道垂直。图 3－2 给出原子轨道 s 与 p 的三种杂化示意图。

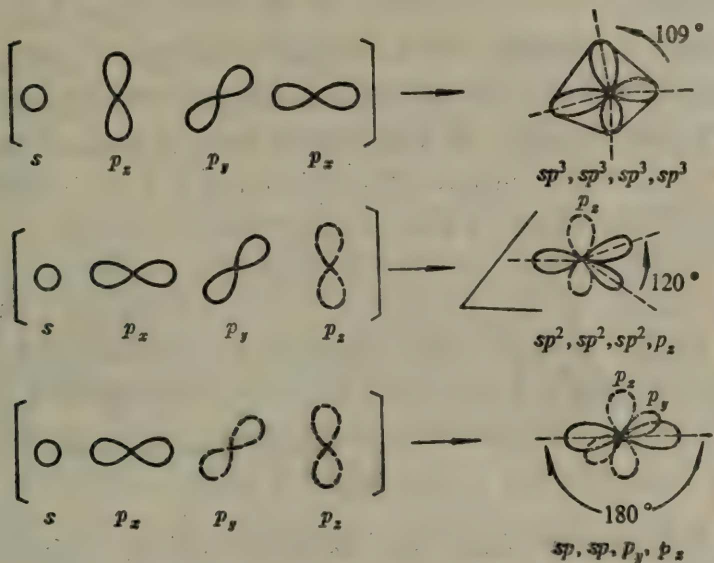


图 3-2 原子轨道 s 与 p 的杂化

有关"杂化"的详尽讨论不是本书要考虑的。但是，必须指出，杂化这个概念仅仅是一种近似，在这里隐含着假定来自同一原子的所有键角都是相等的。然而，这种近似对于一个原子的所有的键在实际上是相等的分子(如甲烷等)才是正确的。

符合杂化轨道定义的数学处理涉及到轨道的线性组合。每一个杂化的 sp^{λ^2}原子轨道可由线性组合 $s+\lambda p$ 来描述(其中 λ 是杂化参数)。如果轨道具有一半 p 成分(sp)，则 $\lambda^2=1$。两个 sp 轨道的轴间夹角为 180 °，而 90 °的键角则相应于纯的 p 成分($\lambda^2=\infty$)。如果原子轨道的 s 成分增加，则会导致键角增大。迄至目前为止，还仅只是研究了 λ 的特殊值，例如 $\lambda^2=3$ 时，则为 sp^3 杂化轨道。

由图 3－3 可见，2 p 轨道能以两种不同的形式相重叠。当轨道以头-头重叠（图3-3A）方式成键时，生成的键对于键轴而言为圆柱形对称，这种键称为 σ 键。杂化轨道也能以这种头-头重叠方式形成 σ 键。当轨道以侧面重叠(图 3－3 B)方式成键时，生成的键对于键轴而言不是圆柱形对称，这种键称为 π 键。

从杂化轨道出发，可以看出，在乙烯分子中当两个 sp^2 杂化碳原子形成一个 σ 键和一个 π 键

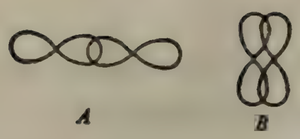


图 3-3　p 轨道的两种重叠方式

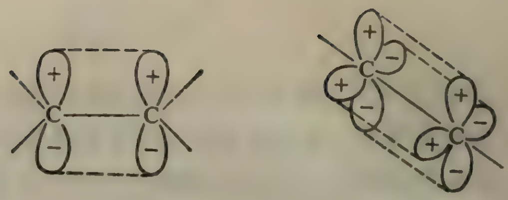


图 3-4

时(图 3－4)，由于 π 键不是圆柱形对称的，所以当沿 σ 键轴旋转时，π 键将发生破裂。这也说明了乙烯分子为平面构型。

氮原子在基态时的电子组态为$(1s)^2(2s)^2(2p_x)^1(2p_y)^1(2p_z)^1$。根据这样的电子组态，可以推测氨分子中氮原子可能有三个正交的 σ 键生成，其孤对电子没有任何最优势的取向。但是这些推测与实验事实是不相吻合的。事实上，氮原子具有四面体成键取向，孤对电子占有着四面体的一角(见图 3－5)。在氨分子中，$R=R'=R''=H$，键角为 107°；在三甲胺分子中，$R=R'=R''=(-CH_3)$，键角为 108°7′。



图 3-5

在氨分子中，氮原子的结构可以通过假设其原子轨道为 sp^3 杂化予以解释。即：

$$(1s)^2(2sp^3)^2(2sp^3)^1(2sp^3)^1(2sp^3)^1$$

其中，有一个杂化轨道$[(2sp^3)^2]$为两个电子所填充，它们起着相当于一个取代基作用的电子对(见图 3－5)。

氧原子在基态时的电子组态为$(1s)^2(2s)^2(2p)^4$。在 8 个电子中有 4 个电子在 $2p$ 轨道，即$(2p_x)^2(2p_y)^1(2p_z)^1$。如果我们假定，在水分子中氧原子上的两对孤对电子占据着两个 sp^3 杂化轨道，余下的两个杂化轨道各包含一个未成对电子，它们可以跟其它原子(氢原子)形成两个化学键。于是，当氧原子与两个氢原子结合形成水分子时，氧原子是利用这两个各包含一个未成对电子的杂化轨道和两个氢原子的 $1s$ 轨道重叠而形成较稳定的化学键，计算的键角为109°28′，跟实验测得的数值 105°相比较，仍有偏差。可以认为，这种偏差的产生是由于氧原子上的孤对电子不参加成键作用，电子云较密集于该原子的周围，这样才能更有利于能量的降低。因此孤对电子所占据的杂化轨道具有较多的 s 轨道成分，其它杂化轨道则含有较多的 p 轨道成分。这种由于有孤对电子的存在而造成的不完全等同的杂化轨道，叫做不等性杂化轨道。或者可以认为，由于孤对电子与孤对电子间的相互排斥作用而使孤对电子彼此离开，从而使两个氢原子靠近。这就解释了为什么水分子的键角是 105°，同时也显示出氧原子的 sp^3 杂化。

跟氧原子同处于第六主族的硫元素，其基态原子参与原子间相互作用时的价态类似于氧原

子的特性(见表 3-2)。但在化合物 R—S—R′ 中,其键角接近于 90°(例如在 H_2S 中为 92°)。在硫的三元取代衍生物中,硫原子成键也是平面型的。

磷原子比硫原子少了一个电子,其基态时的电子组态为 $(1s)^2(2s)^2(2p)^6(3s)^2(3p_x)^1(3p_y)^1(3p_z)^1$,在 M 壳层中的电子排布跟氮原子的相同。因此,与氮原子的情形一样,孤对电子占据着四面体的一个角。

综上所述,碳、氮、氧、硫和磷等构成生物分子的主要元素,按照杂化轨道概念它们往往是以四个杂化轨道参与分子中原子间的相互作用,并决定着共价键的方向性。这对于我们探讨生物分子结构和活性,无疑是有益的。除此而外,氮、氧、氯和硫等元素还因具有较强的电负性,在一些分子中它们可以作为 H^+ 的受体而形成氢键(见 3.2.2)。还须指出,以上所考虑的原子的价态问题,是意味着原子存在于分子中的状态。因为孤立的原子在分子中是不存在的,因此价态概念仅仅是一种近似的概念。

3.1.3.2 离域键

从上面的讨论中可以看出,定域键模型在讨论诸如甲烷、氨、水、乙烯以及其它一些分子的结构时,不仅能够说明化学键的定域特性,还可以定性地给出分子结构中的每一个键的图象。但是,这样的可能性并不是普遍存在的。

定域键模型只是对分子结构的一种近似描述。实际上,分子中的电子并不完全定域。即使象在饱和烃中,电子也是离域于整个分子的。事实已经说明,定域键模型在描述大量的其电子可以离域在两个以上原子的任何分子(如芳香族化合物、核酸碱基等)时就变得不适宜了。典型的芳香族化合物分子是苯(C_6H_6)。苯是平面型分子,它的六个碳原子占据着正六边形的六个角。与这些碳原子的价态相应的原子轨道波函数为:

$$
\begin{aligned}
\phi_1 &= \frac{1}{\sqrt{3}}(\phi_{2s} + \sqrt{2}\,\phi_{2p_x}) \\
\phi_2 &= \frac{1}{\sqrt{6}}(\sqrt{2}\,\phi_{2s} - \phi_{2p_x} + \sqrt{3}\,\phi_{2p_y}) \\
\phi_3 &= \frac{1}{\sqrt{6}}(\sqrt{2}\,\phi_{2s} - \phi_{2p_x} - \sqrt{3}\,\phi_{2p_y}) \\
\phi_4 &= \phi_{2p_z}
\end{aligned}
\tag{3.6}
$$

(3.6)式中,前三式碳原子的 ϕ_1,ϕ_2 和 ϕ_3 相结合并和氢原子的 1s 轨道结合形成分子平面的定域 σ 键,键角为 120°(图 3-6 a)。ϕ_4 则沿伸到分子平面的上方和下方(图 3-6 b)并相结合成 π 键。这个 π 键为六个碳原子所共有,是离域键。如果按简单分子轨道方法,六个碳原子的 ϕ_z(记为 ϕ_{z1}、ϕ_{z2}、ϕ_{z3}、ϕ_{z4}、ϕ_{z5} 和 ϕ_{z6})经线性组合得 π 分子轨道波函数为:

$$
\psi_1 = \frac{1}{\sqrt{6}}(\phi_{z1} + \phi_{z2} + \phi_{z3} + \phi_{z4} + \phi_{z5} + \phi_{z6})
$$

$$
\psi_2 = \frac{1}{2}(\phi_{z2} + \phi_{z3} - \phi_{z5} - \phi_{z6})
$$

$$\psi_3=\frac{1}{2\sqrt{3}}(2\phi_{z1}+\phi_{z2}-\phi_{z3}-2\phi_{z4}-\phi_{z5}+\phi_{z6})$$

$$\psi_4=\frac{1}{2\sqrt{3}}(2\phi_{z1}-\phi_{z2}-\phi_{z3}+2\phi_{z4}-\phi_{z5}-\phi_{z6}) \tag{3.7}$$

$$\psi_5=\frac{1}{2}(\phi_{z2}-\phi_{z3}+\phi_{z5}-\phi_{z6})$$

$$\psi_6=\frac{1}{\sqrt{6}}(\phi_{z1}-\phi_{z2}+\phi_{z3}-\phi_{z4}+\phi_{z5}-\phi_{z6})$$

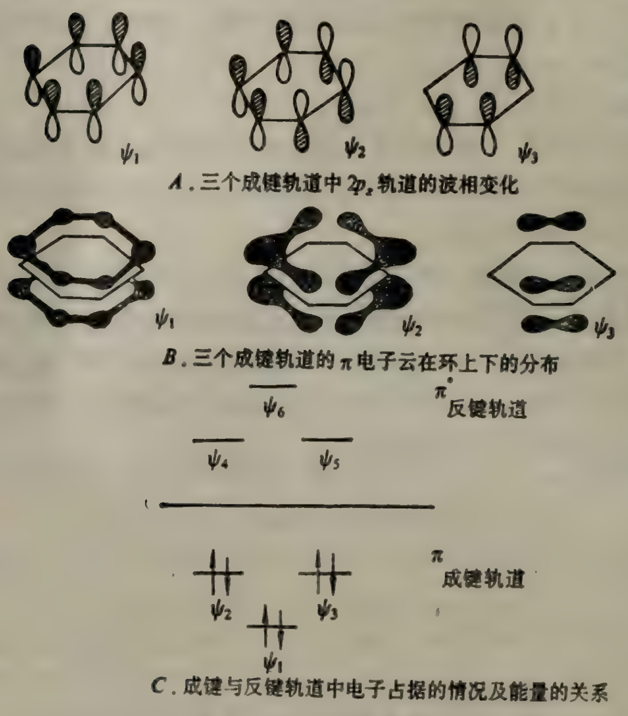


图 3-6 苯分子的成键图示

六个可能的 π 分子轨道中，有三个是成键的(图 3-6 c)。在基态时，占据这三个轨道的六个电子将不再属于某个特定的碳原子，它们可以在形如"汽车的内胎"里运动。由于每一个电子有较多的空间可占据，正是这种非定域的缘故，这些电子的能级都比较低。这个额外的能量(对于苯是 1.8 eV)增加了键的强度。

如果用虚线来描述离域键的话，则苯的结构式如图 3-7 所示。实际上，在苯的所有C═C键中电荷密度都是按同一方式分布的，且原子间的距离也都是相同的(0.14 nm)。但是，在实际书写苯分子的结构时，为简便起见往往是用下述形式来表示的(图 3-8)：

当然，在沿用经典化学结构理论的这种描述方式时，我们假定读者对于离域键的含意已经是了解的了。同时，对于其它芳香族化合物分子的离域键结构也是按这种方式处理的。只不过在

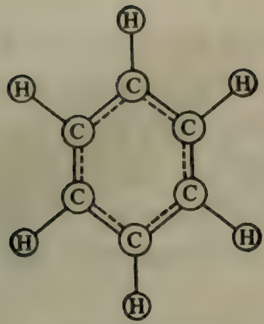
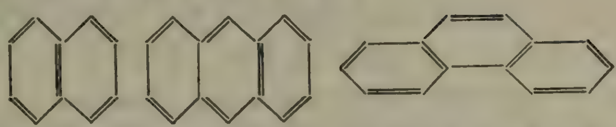


图 3-7 苯的结构式



图 3-8

这个时候，我们还必须考虑到苯核的并接方式，例如：



根据以上讨论，我们可以定义所谓的共轭体系：一般所谓的共轭体系是指不饱和键（双键或叁键）与单键彼此相间所组成的体系。例如：

$CH_2{=}CH-CH{=}CH_2$，$CH_2{=}CH-C\equiv CH$，，

$CH\equiv C-C\equiv H$， 及其它芳香族化合物等。

在生命过程中具有重要生物学作用的共轭体系常常含有氮、氧和硫等杂原子，如组成核酸的嘌呤和嘧啶碱基以及组成蛋白质的色氨酸和组氨酸等。

必须指出，自本世纪二十年代末由 Heitler，London 和 Pauling 等人发展起来的价 键(VB)理论和由 Hund，Mulliken 和 Slater 等人发展起来的分子轨道(MO)理论，已经成为现代化学键理论的重要内容。从图象上看，VB 理论认为分子是由定域键将原子结合在一起而构成的，MO 理论则主张电子的离域化，即认为每个分子轨道都遍及于分子整体。由于 VB 理论是与分子结构经典图象的定域键相联系的，因此一建立便得到迅速发展，但由于计算上的困难曾使其一度停滞，直至六十年代后才又有了新的发展。可是MO 理论在五十年代就有了很大的发展，特别是六十年代其离域化的概念得到了光电子能谱实验的支持，更促进了它的发展。当前，这两种理论都在继续发展。我们在实际处理原子之间和分子之间的相互作用时，应该充分注意到这两种理论的特点及其相互间的联系。

3.2 分子间的相互作用

以上我们讨论了原子间的相互作用，就定域键和离域键作了简述。

在生物分子中，从原子之间的相互作用方式来看，以形成牢固的共价键为一个极端，相互的

排斥作用为另一个极端，表现形式多种多样，作用范围十分广泛。以蛋白质多肽链为例，主链上的共价键决定着分子的线型结构，而弱的相互作用，特别是氨基酸侧链间的相互作用对多肽分子的形态有相当的影响。这样，在生物大分子(如蛋白质和核酸)中既存在着相邻原子间强烈的吸引作用(化学键)，又存在着分子间较弱的吸引作用。这后一种吸引作用是决定生物大分子的沸点、熔点、溶解度、表面张力、粘度等物理化学性质时的主要因素；是决定碱基配对，DNA 双螺旋稳定性，蛋白质的 α-螺旋结构和 β-折迭结构的主要原因之一，在维持生物大分子的构象中起重要作用。同时，也是决定生物分子间专一相互作用的因素。以下我们将讨论生物分子间的这些相互作用。

3.2.1 Van der Waals-London 相互作用

早在 1873 年，Van der Waals 就注意并研究了分子与分子之间存在着的弱的吸引力——分子间的 Van der Waals 吸引力。关于这种力的本质，直到 1930 年 London 提出 Van der Waals 引力的量子理论[4]，人们才对这一作用力的本质有了较深入的了解。

在量子生物学的研究中，讨论分子间相互作用时，按照量子力学的微扰方法所得到的能量分解，可将两分子间的相互作用能表达为[5]：

$$E_{1-2}=E_{el}+E_{pol}+E_{disp}+E_{rep} \tag{3.8}$$

其中 E_{el} 是静电部分，E_{pol} 是极化部分，E_{disp} 是色散部分而 E_{rep} 则为排斥部分。以下分述之。

(1) 静电部分(静电项)

分子 1,2 的静电相互作用能作为一级微扰项而得到

$$E_{el}=\langle\Psi_0^{(1)}\Psi_0^{(2)}|V|\Psi_0^{(1)}\Psi_0^{(2)}\rangle \tag{3.9}$$

式中，$\Psi_0^{(i)}$是指分子 $i(i=1,2)$的无微扰 Hamilton$\hat{H}_0^{(i)}$的基态本征函数，而 V 是类似于双原子分子的Hamilton中相互作用微扰算符的分子间微扰，包含有分子 1 对分子 2 (以及分子 2 对分子1)的原子核间排斥，电子间排斥以及核与电子间的吸引等。

为了处理(3.9)式的矩阵元，按照静电理论，可以把电荷分布于一个小区域 τ 内的电荷体系在远处所产生的势表示成很多等效电极子在远处所产生的电势的叠加。于是可以将势 V 代之以适于各分子的单中心多极展开，或者等价地在单个分子中心进行每个分子电荷密度的多极展开。这种多极展开仅能准确地表达集中在原点并包含全部电荷分布的球之外电荷分布的势。因此，当分子不是球形时就有所不足[5]。

为此，代替 V 的简化，我们取一种仅与所涉及到的分子电荷密度有关的简化。实际上在 MO-LCAO 范围内，电荷密度的电子部分最后简化为单原子和双原子基元分布 $\phi_\alpha(\boldsymbol{r})\phi_\beta(\boldsymbol{r})$的和($\phi_\alpha$ 等是所用基组的原子轨道)。这些分布有三种类型：单中心"对角"分布($|\phi_\alpha|^2$)和非对角分布($\phi_\alpha^*\phi_{\alpha'}$)以及双中心分布($\phi_\alpha^*\phi_\beta$)(这种分布仅出现在原子间的重叠积分不忽略的情形，例如在 ab initio范围内)。于是每一个小线度的原子或键分布可以足够准确地用在各自恰当中心的多极展开来表达，结果总的分子电荷分布将表现为与不同中心有关的若干个多极展开的和(也称作多中心多级展开)。这种展开由于定义明确尤其适合于通常的分子形状。有关多中心多极展开的应用正在进一步发展[6-7]。

关于单中心分布，有下列特性：如果多极展开是对于这种分布的(本来)原子中心进行的，则仅有有限数目的非零多极矩。更确切些，如果所包含的两个原子轨道 $\phi_\alpha(\boldsymbol{r})\phi_{\alpha'}(\boldsymbol{r})$ 有角量子数 l_α 和 $l_{\alpha'}$，那么非零的多极矩仅只有 n 级($|l_\alpha-l_{\alpha'}|\leqslant n\leqslant|l_\alpha+l_{\alpha'}|$)。这一性质是球谐函数加法定理的直接结果(本身也显示了角动量的加和定理)。

当原子间重叠积分不予忽略时，我们必须考虑双中心分布 $\phi_\alpha{}^*\phi_\beta$，这种分布的电荷是重叠积为 $S_{\alpha\beta}=\int\phi_\alpha{}^*(\boldsymbol{r})\phi_\beta(\boldsymbol{r})\mathrm{d}\boldsymbol{r}$。现在多极展开包含有无限个非零项，通常我们只考虑电荷(零极矩)、偶极矩和四极矩：

$$
\begin{aligned}
q&=\int\phi_\alpha{}^*\phi_\beta\mathrm{d}\boldsymbol{r}\text{，电荷}\\
\boldsymbol{\mu}&=\int\phi_\alpha{}^*\boldsymbol{r}\phi_\beta\mathrm{d}\boldsymbol{r}\text{，偶极矩}\\
\boldsymbol{Q}&=\int\phi_\alpha{}^*[\boldsymbol{r}\otimes\boldsymbol{r}]\phi_\beta\mathrm{d}\boldsymbol{r}\text{，四极矩}
\end{aligned}
\tag{3.10}
$$

四极矩是一个二阶对称张量，在静电相互作用计算中，我们只用到轴向四极矩[8]。如果用 $\boldsymbol{a}$ 表示给定轴的单位向量而 θ_a 表示相应的四极矩矩阵元，则

$$\boldsymbol{Q}_{\text{axial}}=\theta_a(\boldsymbol{a}\otimes\boldsymbol{a}) \tag{3.11}$$

用了这些电荷分布，多极-多极相互作用能用下列标准公式计算。在这些公式中，$\boldsymbol{r}$ 表示连接两个多极子的向量，而 $\boldsymbol{u}=\boldsymbol{r}/r$ 表示相应的单位向量。

单极-单极能

$$E=q_iq_j/r \tag{3.12}$$

单极-偶极能

$$E=-\boldsymbol{\mu}_j\boldsymbol{\varepsilon}q_i \tag{3.13a}$$

其中

$$\boldsymbol{\varepsilon}q_i=q_i\boldsymbol{u}/r^2=q_i\boldsymbol{r}/r^3 \tag{3.13b}$$

表示由电荷 q_i 在定位 $\boldsymbol{\mu}_j$ 的 j 点所产生的电场，因此 $\boldsymbol{r}=\boldsymbol{r}_{ij}=\boldsymbol{r}_j-\boldsymbol{r}_i$。

偶极-偶极能

$$
\begin{aligned}
E&=-\boldsymbol{\mu}_j\boldsymbol{\varepsilon}_{\mu i}\\
&=\frac{1}{r^3}[\boldsymbol{\mu}_i\boldsymbol{\mu}_j-3(\boldsymbol{\mu}_i\boldsymbol{u})(\boldsymbol{\mu}_j\boldsymbol{u})]
\end{aligned}
\tag{3.14a}
$$

其中

$$\boldsymbol{\varepsilon}_{\mu i}=-\frac{1}{r^3}[\boldsymbol{\mu}_i-3(\boldsymbol{\mu}_i\boldsymbol{u})\boldsymbol{u}] \tag{3.14b}$$

表示由偶极 $\boldsymbol{\mu}_i$ 在偶极 $\boldsymbol{\mu}_j$ 所在位置产生的电场。

单极-四极能(轴向四极：$\boldsymbol{a},\theta_a$)

$$E=q\frac{\theta_a}{2r^3}[3(\boldsymbol{a}\cdot\boldsymbol{u})^2-1] \tag{3.15}$$

偶极-四极能(轴向四极：$\boldsymbol{a},\theta_a$)

$$
\begin{aligned}
E&=-\boldsymbol{\mu}\boldsymbol{\varepsilon}_\theta\\
&=\frac{3\theta_a}{2r^4}\left\{2(\boldsymbol{a}\cdot\boldsymbol{u})(\boldsymbol{a}\cdot\boldsymbol{\mu})+[1-5(\boldsymbol{a}\cdot\boldsymbol{u})^2](\boldsymbol{\mu}\boldsymbol{u})\right\}
\end{aligned}
\tag{3.16a}
$$

其中

$$\boldsymbol{\varepsilon}_\theta = -\frac{3\theta_s}{2r^4}\left\{2(\boldsymbol{a}\cdot\boldsymbol{u})\boldsymbol{a}+[1-5(\boldsymbol{a}\cdot\boldsymbol{u})^2]\boldsymbol{u}\right\} \tag{3.16b}$$

表示电四极在电偶极所在位置产生的电场。

四极-四极能：两个轴向四极($\boldsymbol{a}_1,\theta_1$)和($\boldsymbol{a}_2,\theta_2$)的相互作用能由下式给出：

$$E=\frac{3\theta_1\theta_2}{4r^5}[1-5(\boldsymbol{a}_1\cdot\boldsymbol{u})^2+(\boldsymbol{a}_2\cdot\boldsymbol{u})^2+2(\boldsymbol{a}_1\cdot\boldsymbol{a}_2)^2 -20(\boldsymbol{a}_1\cdot\boldsymbol{u})(\boldsymbol{a}_2\cdot\boldsymbol{u})(\boldsymbol{a}_1\cdot\boldsymbol{a}_2)+35(\boldsymbol{a}_1\cdot\boldsymbol{u})^2(\boldsymbol{a}_2\cdot\boldsymbol{u})^2] \tag{3.17}$$

(2) 极化部分(极化项)

分子 1 被分子 2 所极化的极化能按下式计算

$$E_{\rm pol}^{(1)}=-\frac{1}{2}\sum_i \alpha_i \varepsilon_i^2 \tag{3.18}$$

式中的 ε_i 是分子 2 的多极子在分子 1 的各个中心 i(原子或键的中点)所产生的电场，α_i 则是根据i 是一个原子或是一个化学键的中点而定义的原子或键的极化率。

电场 ε_i 的计算逻辑上由静电项中所用到的多极展开而引伸出，α_i 的定义要与多极子的定位一致。中心在分子 2 的j点的电荷 q_j，偶极 $\boldsymbol{\mu}_j$ 或轴四极 θ_j 在分子 1 的点 i 所产生的电场按公式(3.13b)、(3.14b)、(3.15)计算，由于指标 i、j 所表示的意义不同，在这里公式定为：

$$\left.\begin{aligned}
\boldsymbol{\varepsilon}_{q_j}(i)&=\frac{q_j}{r_{ji}'^3}\boldsymbol{r}_{ji}\\
\boldsymbol{\varepsilon}_{\mu_j}(i)&=-\frac{1}{r_{ij}'^3}\left(\boldsymbol{\mu}_j-\frac{3(\boldsymbol{\mu}_j\cdot\boldsymbol{r}_{ji})\boldsymbol{r}_{ji}}{r_{ji}'^2}\right)\\
\boldsymbol{\varepsilon}_{\theta_j}(i)&=-\frac{3\theta_j}{2r_{ij}'^4}\left\{2(\boldsymbol{a}\cdot\boldsymbol{u})\boldsymbol{a}+[1-5(\boldsymbol{a}\cdot\boldsymbol{u})^2]\boldsymbol{u}\right\}
\end{aligned}\right\} \tag{3.19}$$

其中 $\boldsymbol{r}_{ji}$ 是连结 i 和 j 的向量，r_{ij} 是它的长度，q_j、$\boldsymbol{\mu}_j$ 和 θ_j 是位于j点的电荷，偶极和轴向四极的强度，$\boldsymbol{a}$ 是所选轴的方向，$\boldsymbol{u}$ 是沿着 ji 方向连结 i 和 j 的单位向量。为了减小极化能的绝对 值，实际上也是为了避免当 r_{ij} 趋于零时极化能无限大，必须以 r_{ij}' 代替 r_{ij}

$$r_{ij}'=r_{ij}+0.5D(R_i^{\rm W}+R_j^{\rm W}) \tag{3.20}$$

D的值视(3.18)式中按原子或键中点为中心来进行加和而分别取作 0.10 或 0.45。$R_i^{\rm W}$ 和 $R_j^{\rm W}$分别是原子i、j的Van der Waals半径(如果 i 或 j 是一个键的中点，则取该键两端原子的Van der Waals 半径的算术平均值)。

极化率 α_i 的取值是由键极化率张量(主轴之一沿键方向)的平均分量 $\alpha_{\rm L}$ 和垂直分 量 $\alpha_{\rm T}$ 而得到，取其平均

$$\bar{\alpha}=\frac{1}{3}(\alpha_{\rm L}+2\alpha_{\rm T}) \tag{3.21}$$

键 I —J 的原子极化率 α_I 和 α_J 以及键极化率 α_{IJ} 定义为

$$\left.\begin{aligned}
\alpha_I&=\bar{\alpha}\frac{n_I}{N_I}\sum_j^{N_I}\left(n_{IJ}+\frac{n_I}{N_I}+\frac{n_J}{N_J}\right)^{-1}\\
\alpha_J&=\bar{\alpha}\frac{n_J}{N_J}\sum_i^{N_J}\left(n_{IJ}+\frac{n_I}{N_I}+\frac{n_J}{N_J}\right)^{-1}\\
\alpha_{IJ}&=\bar{\alpha}n_{IJ}\left(n_{IJ}+\frac{n_I}{N_I}+\frac{n_J}{N_J}\right)^{-1}
\end{aligned}\right\} \tag{3.22}$$

式(3.22)中，N_I 和 N_J 分别是始于 i、j 的键数，n_I 和 n_J 分别是原子 i 和 j 的孤对电子数，n_{IJ}是键 $I-J$ 之上的共价电子数。这个公式相当于把各个平均键极化率 $\bar{\alpha}$ 按照由"成键的"和"不成键的"(孤对电子)的价电子数计算而得到的"权"，均分为两个原子分量和一个键分量。于是属于原子 i 的所有 α_I 求和便得到 α_i。

表 3-2 给出了用于上述计算的某些键极化率[9]；表 3-3 则给出根据同系物分子的实验极化率而得到的最适宜静电原子极化率，还有电离势[10]，它们常用于生物分子间相互作用的计算。

表 3-2 键极化率

键	α_L ($\times 10^{-3}$nm^3)	α_T ($\times 10^{-3}$nm^3)
C—C(脂肪)	1.88	0.02
C—C(芳香)	2.25	0.48
C═C	2.86	1.06
C≡C	3.54	1.27
C—H(脂肪)	0.79	0.58
C—Cl	3.67	2.08
C—Br	5.04	2.88
C═O(羰基)	1.99	0.75
C═O(CO_2)	2.05	0.96
C═S (CS_2)	7.57	2.77
C≡N(HCN)	3.1	1.4
N—H(NH_3)	0.58	0.84
S—H(H_2S)	2.30	1.72

(3) 色散和排斥部分

分子间相互作用的色散-排斥能部分，在早期 Kitagorodsky[11]公式的基础上，又有了改进[12]。

$$E_{dis-rep}=\sum_i^{(1)}\sum_j^{(2)}E(i,j) \tag{3.23}$$

i 指分子 1 中的原子，j 指分子 2 中的原子，每一对原子-原子贡献是色散和排斥项的和

$$E(i,j)=K_iK_j\left[-\frac{A}{Z^6}+C\cdot\exp(-\alpha z)\right] \tag{3.24}$$

其中

$$Z=R_{ij}/R^0_{ij} \tag{3.25}$$

$$R^0_{ij}=\sqrt{(2R^W_i)(2R^W_j)} \tag{3.26}$$

R_{ij} 是原子 i、j 间的距离，R^W_i、R^W_j 是原子 i、j 的 Van der Waals 半径(一般取自文献[13])。参数 α、A、C 与原子类别 i、j 无关，参数 K_i 则与原子类别有关。

考虑到电子布居的变化对排斥贡献会有影响[12]，故对每一原子-原子排斥项引入乘因子

$$(1-Q_i/N^V_i)(1-Q_j/N^V_j)$$

N^V_i 是原子 i 的价电子数，Q_i 是它的净电荷，这样

$$E_{rep}(i,j)=(1-Q_i/N^V_i)(1-Q_j/N^V_j)K_iK_jC\cdot\exp(-\alpha Z) \tag{3.27}$$

表 3-3 最适宜原子静电极化率和电离势

原 子	价 态①	$\alpha_i(\times10^{-3}nm^3)$	I_i(eV)
C	tetetete(C1)	1.064	14.57
	trtrtrπ(C2)②	1.332	11.22
	trtrtrπ(C3)	1.230	11.22
	trtrtrπ(C4)	1.529	11.22
	didi$\pi\pi$(C5)	1.279	11.24
N	te^{2}tetete(N1)	1.094	14.31
	trtrtrπ^2(N2)	1.090	12.25
	tr^{2}trtrπ(N3)	1.030	14.51
	di^{2}di$\pi\pi$(N4)	0.852	14.47
O	te^{2}te^{2}tete(O1)	0.664	18.40
	tr^{2}tr^{2}trπ(O2)	0.460	17.25
	tr^{2}trtrπ^2(O3)	0.422	14.97
	te^{2}te^{2}te^{2}te(O4)	1.791	6.31
H	σ(H)	0.386	13.61
P	teteteteπ(P)	1.743	12.09

① 轨道类型：te＝四面体，，tr＝三角，di＝对角，σ，π。
② 各价态按括号中符号标记，其中，C2——脂肪烃，C3——芳香烃，，C4——缩合烃。

上述校正多数情形很小，当 H 原子与 O 或 N 原子键合时则要注意，因为此时有不可忽略的正电荷，且 $N_H^V=1$。

对于有氢键形成的原子-原子对，由于电荷转移的贡献，(3.27)式的校正仍然不足。为了表明这一点，在短距离（对于氢与重原子相互作用）采用一个按下述方法定义的降低了的排斥项：选择两个距离 $R_m<R_M$；对于 $R>R_M$，用通常的参数 A,C,α；对于 $R<R_m$，用调整后的参数 A'，C'，α'；对 $R_m<R<R_M$，则用内插值。详见文献[12]。

与讨论分子结构的量子化学计算方法类似，在分子间相互作用能的计算中，选用不同的近似等级，相互作用势函数就不同。再者，对公式中参数的选择，原则上以能重现分子晶体的某些实验数据为好。因此对不同的化合物系，拟合的参数会有差异。尽管如此，分子间的 Van der Waals-London 相互作用确是真实存在于分子或原子之间的一种作用力。其作用能虽比化学键键能小一至二个数量级，也有一定的范围，但对于决定分子（特别是生物分子）的结构和性质仍是不可忽视的因素；对于量子生物学的理论计算有着其不寻常的作用，根据作者等的经验，尤其在讨论生物系统中分子间的相互作用（如专一识别相互作用）时，这种作用力的理论计算就显得更加重要。因此，在这一节里，我们不惜用一定的篇幅讨论这一问题。作为理论研究的例子，早在二十多年前，为了了解和预言溶液中核酸的性质，需要弄清是哪些力在支配着多核苷酸的分子构型。然而，要从实验上来判断溶液中双股螺旋多核苷酸相对于两单链自由盘绕型的稳定性时，其中各种力的贡献有何不同？是非常困难的。在理论研究方面，Do Voe 和 Tinoco[14] 计算了

DNA 的杂环碱基对溶液中上述两种构型的自由能的主要贡献。他们采用碱基的基团偶极矩和极化率，通过计算表明，在螺旋的碱基之间存在着较大的相互作用。虽然计算的近似等级低，却得出了定性的结论，概括起来有：

(1) 静电力和色散力可以影响 DNA 的碱基序列，由此减少了序列所可能包含的信息量；

(2) GC 和 AT 多聚物的相对解链温度与溶剂有关；

(3) 在极少溶剂存在下，平行碱基间大的色散力能引起碱基按一对一的多核苷酸链堆积成有序的排列。

二十年来，随着高速电子计算机的使用，有关分子电子壳层计算方法的精心设计以及比级高级的数学方法的运用，使分子间相互作用的理论计算得以迅速发展。与此同时，对 DNA 碱基间相互作用的研究也越来越深入，其中 Langlet 和 Claveric[8]等对氢键加合物和堆积加合物作了较好的理论计算。我国科学工作者江寿平、乐树云等[15]对 RNA 12 中互补碱基共平面配对的相互作用能作了计算，进而作了较详尽的分析(见 3.3.3)。我们也计算了含甲基化碱基的堆积相互作用能(表 3-4)。

表 3-4 碱基堆积相互作用能*(kcal/mol)

	静电能	极化能	色散能	排斥能	总能量
C·C	−0.059	−0.209	−3.492	0.354	−3.406
m^5C·C	0.063	−0.279	−6.951	2.417	−4.750
C·m^5C	−0.035	−0.218	−3.572	0.355	−3.470
G·C	−0.676	−0.176	−2.227	0.165	−2.914
G·m^5C	−0.721	−0.183	−2.270	0.165	−3.009
C·G	−0.963	−0.366	−9.528	2.035	−8.822
m^5C·G	−1.099	−0.432	−13.184	4.169	−10.546
G·G	1.115	−0.307	−7.631	1.439	−5.384
C·A	0.132	−0.120	−8.590	1.917	−6.662
m^5C·A	0.170	−0.178	−12.755	4.508	−8.255
A·m^5C	2.448	−0.714	−3.464	0.279	−1.451
T·m^5C	−1.631	−0.328	−7.612	2.360	−7.211
m^5C·T	−0.192	−0.168	−6.935	2.223	−5.072
T·A	−0.347	−0.209	−11.386	3.992	−7.950
A·G	−3.710	−0.780	−8.685	1.677	−11.498
T·G	1.773	−0.469	−13.956	4.770	−7.882

* m^5C·C表示 ↑ $\frac{C}{m^5C}$

3.2.2 分子间的氢键相互作用

上一节，我们讨论了分子间的 Van der Waals-London 相互作用。现在，我们来讨论生命系统中分子的另一类重要的相互作用——氢键。

氢键相互作用是质子给予体 X—H 和质子接受体 Y 之间一种特殊类型的相互作用，是一种在流动的氢原子和电负性很强的杂原子(F、O、N、Cl 和 S)之间起作用的键：

$$X—H\cdots YR$$

其中 X、Y 代表 F、O、N、Cl 和 S 等电负性大而半径小的原子。

例如，在吡啶分子中，氮原子有两个自旋反向的外层电子(即孤对电子)没有参与直接的相互作用。这种“游离的”或“没有分开的”电子对可以吸引质子而形成键(当其中一个吡啶分子转变成离子状态时)(图 3-9)。这时，一个吡啶分子的氮原子便与另一个处于离子状态的吡啶 $(C_5NH_6)^+$ 的氮杂原子结合生成分子间的复合物。在复合物中由于氢键结合的影响，质子既可处在靠近左边的氮原子的位置，也可以处于靠近右边的氮原子的位置。因而质子的势能应当随着至左边或至右边氮原子的距离形成具有两个最小值的曲线(当距离一定时，如 0.31 nm)。Rein 等[16]对这样的曲线作了量子力学计算(见图 3-10)。

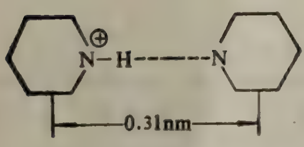


图 3-9 吡啶分子间的氢键

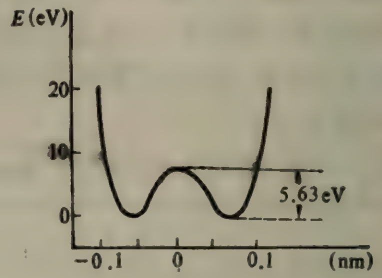


图 3-10 在两个吡啶分子的氮原子间随位置改变的质子势能曲线

氢键可分为分子间氢键和分子内氢键两大类。一个分子的质子给予体 X—H 与另一分子的质子接受体 Y 相结合而成的氢键叫做分子间氢键。而一个分子的质子给予体 X—H 与其内部的质子接受体 Y 相结合而成的氢键叫做分子内氢键。

在氢键中，最通常的质子给予体有 OH、NH，而 SH 则是很弱的质子给予体。质子接受体 Y 均有未成键的 p 电子或 π 电子，通常有：

OH，OR，C═O，COOR，环氧化物；

NH_2，NH—R，NR_2，N(芳香氮)，NO_2；

SR，S═O，P→O；

F、Cl、Br、I；

C═C，C═C═C，C═C(芳键)；

C═N，C≡N，N═N 等。

氢键的键能是指发生下列过程所需要的能量：

$$X—H\cdots YR \rightarrow X—H + Y—R$$

氢键比共价键弱，一般比 Van der Waals 力强。在生物系统中，氢键键能在通常情况下是 4—8 kcal/mol，其键长为 0.25—0.32 nm，键角是 180°—125°。为了与有氢参与的共价键相区别，氢键用虚线表示。

我们知道，通过氢键相互作用形成的最简单的分子间复合物是 $(HF)_2$。这种 HF 的二聚体具

有线型结构，氟原子间的距离为 0.279 nm。在 HF 极性分子中，原子间的距离等于 0.092 nm，形成复合物时释放出大约 0.26 eV 的能量。通过氢键相互作用形成的水的二聚体——$(H_2O)_2$，其键能为 0.2 eV。这一能量大约等于 20 个 OH 共价键的能量。在复合物中，两个氧原子间的距离约等于 0.276 nm，氧原子的 Van der Waals 半径为 0.152 nm。Hankins 等[17]对形成复合物时氢原子电子密度的改变作了计算。他们的计算说明，当形成复合物时反应分子中所有原子周围的电子密度分布都发生了变化(见图 3-11)。

以上仅对氢键作了一般的叙述，考虑到它在生命系统分子间相互作用中的重要性，下面我们通过两个问题的讨论来进一步说明氢键的性质和作用。

1) DNA 碱基间的氢键相互作用

我们知道，在生命系统中氢键决定着对所有生命过程具有根本意义的核酸和蛋白质的二级结构。通常认为，通过氢键相互作用，DNA 分子的碱基才得以配对，并导致双螺旋形成稳定的结构。同时，氢键对于复制和转录过程也起着重要的作用。DNA 的自身复制，遗传信息由DNA→RNA→蛋白质的传递都有赖于氢键的断裂和生成。

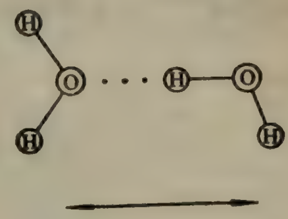


图 3-11 在通过氢键形成的二聚体中，原子附近电子密度的变化

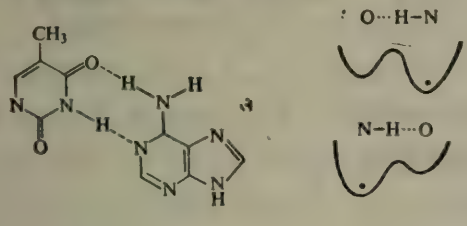


图 3-12 含氮碱基间的氢键 *A*.A-T 碱基对，*B*.氢键的势能曲线

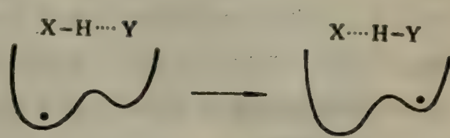


图 3-13

图 3-12 表示，腺嘌呤(A)和胸腺嘧啶(T)在形成 DNA 分子双螺旋结构时，两个氢键相互作用的影响。从图中可以看到，在两个氢键中质子势能曲线的极小值位置的相互关系。Lőwdin提出，每一个氢键体系 X—H···Y 中存在着一个双势阱(图 3-12 *B*)，由于质子的隧道效应，质子从给予体 X-H向接受体 Y 的迁移几率不是零，因而图 3-12 中的体系可以变成另一种体系，即由 X—H···Y 变成 X···H—Y (图 3-13)。譬如在 DNA 中的 A-T(或 G-C) 碱基对的第一个氢键(上面的一个)，通过隧道效应，由氨基—酮型氢键变成亚氨基—烯醇型氢键，结果便产生互变异构碱基对：

$$A\text{-}T \longrightarrow A^*\text{-}T^*$$
$$G\text{-}C \longrightarrow G^*\text{-}C^*$$

从而导致错误复制和突变(见第七章)。

作者曾运用(3.28)式计算了 G-C 碱基对中第一个氢键的质子沿氢键移动时的势能变化，结果如图 3-14 所示。

$$U = \sum_{i=1}^{n}{}' \frac{q_i}{r_{ix}} \tag{3.28}$$

第一个氢键体系 N—H···O（从 0.100 nm—0.184 nm）中不存在双势阱，一个单势阱处在 x 等于 0.130nm 处，且右边的 U高于左边。在此情况下，看来质子向右侧移动的几率是很小的。

有人用 π 电子的 SCFMO 法计算了 G-C 对的三个氢键中中间一个氢键的质子转移时的势能变化[16,18]。计算结果如图 3-15 所示。

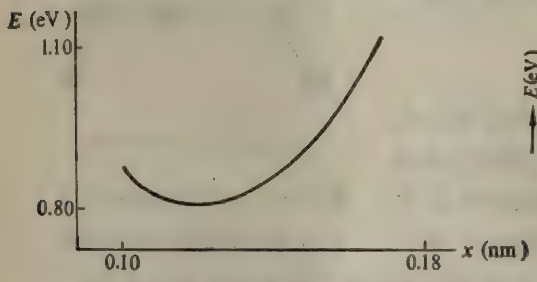


图 3-14 G-C 碱基对第一个氢键(N—H···O)中，因 H 的移动而得到的势能曲线(横坐标是从胞嘧啶的 H 位置起算的距离)

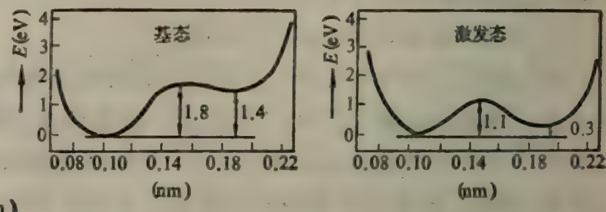


图 3-15 基态、激发态中 G-C 对的氢键(中间的 N—H···N)的质子的势能曲线

在基态，左边的势阱较深，表明质子向右迁移的几率极小。反之，在激发态，左右势阱的深度几乎相同，可以认为质子向右边迁移的几率有显著增加。

关于碱基间的氢键相互作用，读者可参见 4.2.2,4.2.4,4.9.3 和 3.3.3。

2）水中的氢键

我们知道，水是一切生物体的基本组分之一，没有水就没有生命。水是在细胞中进行生物化学反应的基本环境，它参与形成血液和淋巴液。生物聚合物的功能和许多超分子构造都是以同水发生相互作用为前提的。

人们已经研究过水的某些性质，但对于液态水的结构仍旧是不太清楚的。已经知道的是：每个水分子具有较大的电矩。由于氧原子有高的电负性，因而水分子可以跟一个、两个、三个乃至四个其它的水分子生成氢键。在液态水中基本上保持了四面体结构。而分子间络合物的组成和结构则与水的结构有关。

在正常压力和温度在零度以下时，水具有最整齐的晶体结构(冰)，如图 3-16 所示。在冰结构中，每个中心氧原子按正四面体方位与另外四个氧原子连结，氧原子间的距离为 0.276nm。每个水分子通过四个氢键跟其它水分子结合。在这种情况下，OH 键间的键角接近于“四面体”的数值，为109.1°。在游离水分子中，OH 键间的键角等于 105°。

冰的结构类似于金刚石结构，但在金刚石中碳原子间是化学力在起作用。冰的晶体属于分子晶体，在晶体中分子基本上保持着自己的个性和相互间的氢键。在分子间相互作用的全部能量(11.64kcal/mol)中，氢键大约贡献 69%。由于氢键的作用才使水的熔点(0°C)和沸点(100°C)

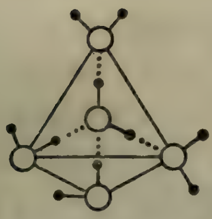


图 3-16

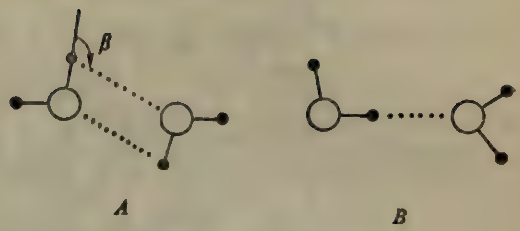


图 3-17 二聚体的氢键
A. β=110°, B. "直线型"氢键(β=0°)

有别于在分子间仅存在着 Van der Waals 力作用的其它液体分子的熔点和沸点。

水有线型的和环状的二聚体以及其它含有 3、4、5、6 和更多分子的络合物。实际上，环状物中分子的数量取决于 OH 键和氢键间的夹角 β。对于二聚体，β 等于 110°，在五环结构中，β 近似于 10°。但在六环和六方晶体结构中 β 角接近于零度("直线型"氢键)，见图 3-17。

自从对水的二聚体第一次运用 ab initio 法作理论计算以来，对一些氢键系统先后作了许多量子化学计算，并且在复合物的氢键能量与分子成分相对方向的预言方面给出了比较可靠的结果。Yamabe 等人[19] 提出了一种能量分解方法：他们把氢键的能量分解成四个部分，即静电能，交换排斥能，极化能和电荷转移或离域能，并在此基础上进行了电子密度分布分析。能量分解的着眼点是整个体系上的能量变化，而电子密度分布分析强调的则是相互作用的那一部分，它经历了电子密度的较大重新分布。对水的二聚体的处理(见图 3-18)表明：(1) 在分子间的区域内，交换相互作用提高了反键合密度，这和交换相互作用的负值(排斥)是一致的。(2) 在质子给予体的水分子中沿着 O—H…O线的氧轨道，氢键质子丢失了相当多的电子密度。而在质子接受体的分子中，电荷重新分布却几乎只归因于极化相互作用。这个效应促使每一对孤对电子和键的极化，它的结果却通过接受体分子的电荷而更迭。极化的如此重要作用，反映在偶极矩的大的变化上。由于质子末端的大的极化，使它的方向并不和 O…H—O轴相一致。(3) 电荷转移的作用从质子接受体移动电荷至质子给予体上。

以上我们列举了 Yamabe 和 Morokuma 等关于氢键分子轨道研究的部分工作。近几年来，有关生物大分子(核酸和蛋白质)在水中，以及药物作用过程中氢键的性质和作用的量子力学计算发展较快，在以后的一些章节中结合具体问题还将继续讨论氢键相互作用。

3.2.3 电荷迁移相互作用

在生物系统中，生物分子还可通过电子给予体分子(D)-电子接受体分子(A)的相互作用而形成"电荷迁移络合物"。电荷迁移是生物体系的重要作用方式和传能方式之一。很多生物体内的反应以及生物化学合成过程都是通过这种相互作用而进行的。电荷迁移在药物与受体的相互作用中也是极为常见的。

电荷迁移络合物可以在离子间形成，也可以在分子间形成。

1) 电子给予体-电子接受体

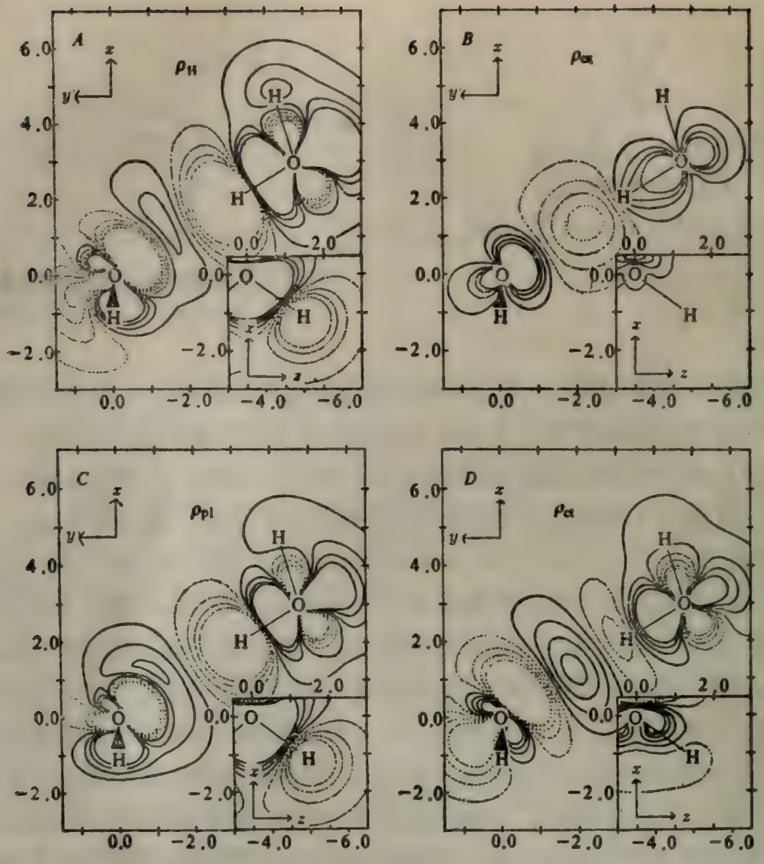


图 3-18　水二聚体中电子密度变化和它的成分，
实线表示密度增加，虚线表示密度减少

电子易于流动的体系(如共轭分子或含有孤对电子的分子)，才能产生电荷迁移。而实现电荷迁移，则要求给予体(D)的 HOMO 和接受体(A)的LUMO 必须是空间重 叠，要 有相同的对称性，即前线轨道的对称性匹配。

在生物体内，生物大分子内部的各种基团之间通过电荷迁移而形成络合物时，可以分别作为电子给予体，电子接受体或者既可作为电子给予体又可作为电子接受体。以组成蛋白质的氨基酸为例，胱氨酸、蛋氨酸和酪氨酸(只限于芳香环)一般情况下是电子给予体；而半胱氨酸，或许还有精氨酸和赖氨酸是电子接受体；组氨酸、丝氨酸、苏氨酸、酪氨酸(只限于羟基)、苯丙氨酸(只限于芳香环)和色氨酸等既可作为电子给予体也可作为电子接受体。从芳香族氨基酸的电子特性来看，它们的给电子能力不如接受电子的能力。一般地讲，它们都是不良的电子给予体(色氨酸在四种芳香氨基酸中是最佳的电子给予体。在组成核酸的碱 基 中，鸟嘌呤表现出良好的 π 电子给予体性质，而尿嘧啶则是不良的电子给予体。

一个分子在参与形成“电荷迁移络合物”时，它是作为给予体或是接受体不仅跟本身的性质有关，而且也跟配偶的性质有关。电荷迁移表明组成络合物的一个分子 D (电子给予体)的电

子可以享用另一个分子A(电子接受体)的轨道[20]。一个电子的迁移不一定涉及分子的重排。从图3-19来看，可以将电荷迁移表示为一个电子从给予体分子(D)的 HOMO 向接受体分子(A)的LUMO 的单纯转移，而不发生任何进一步的重排。

由图 3-19 可以明显地看出，为了使电子能从 D 的一个 HOMO 转移到 A 的 LUMO上，这二个轨道必须是空间重叠，才会出现最大的相互作用。这就要求二个分子必须非常靠近，因此密切配合和可接近性就成了决定性因素。

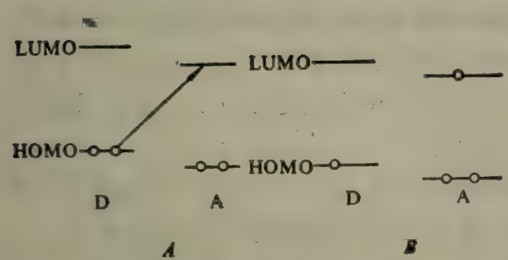


图 3-19 电荷迁移前(A)后(B)，二分子基态与第一激发态

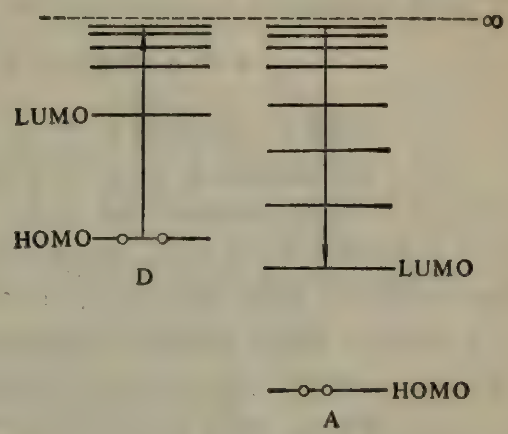


图 3-20

图 3-19 还表明，能级图上二个轨道的相对位置，亦即给予体 D 的电离势(IP)和接受体 A 的电子亲合势(EA)必定是起主导的因素。由图看，应区别两种极端情况：

(1)，D 的HOMO 比 A 的 LUMO 低得多，如图所示，因而使电子由前者上升到后者需要相当大的能量，即要求$IP \gg EA$；(2)，A 的LUMO 比 D 的 HOMO 低，因而电子不需要外加能量就能发生迁移，即 $EA > IP$，如图 3-20 所示。

电荷迁移相互作用的势能与给予体的电离势和接受体的电子亲合势的差值成正比。亦即，从给予体移去电子或把电子加给接受体越容易，则这种相互作用的可能性就越大。

2）电荷迁移相互作用的量子理论

当分子间的作用力达到一定的程度，但还没有形成共价键时，生成的是分子间络合物，且它的吸收光谱与原化合物不同。这时，我们就不能用 Van der Waals-London 相互作用等作用力来描述，而需要用电荷迁移相互作用来予以讨论。

Mulliken 首先对电荷迁移络合物进行了量子力学处理[21]。他证明了这些分子络合物的稳定性与它们的特征电子吸收带两者，都可以和它们的给予体-接受体性质联系起来。Mulliken 指出，电荷迁移络合物(D·A)的状态由 D 与 A 相互作用的非键态(D,A)和配键态($D^+—A^-$)的线性组合而成。用量子力学符号表示，线性组合过程可写成如下形式：

$$\psi(D \cdot A) = a\psi_0(D,A) + b\psi_1(D^+—A^-) \qquad (3.29)$$

式(3.29)中非键波函数 ψ_0 指的是 D 与 A 在相互作用时，只通过 Van der Waals-London 相互作用等力；而配键波函数 ψ_1 则描述给予体的一个电子完全转移给接受体。式中，a 和 b

是组合系数，$a \gg b$，因第二项在基态时不能最大限度地形成络合物。如果吸收足够的能量，则分子间发生一个电子迁移，此时式(3.29)中的第二项占优势，同时产生特殊的电子迁移谱而形成电荷迁移络合物。

非键状态与配键状态(电荷迁移状态)间的相互作用取决于 D 的轨道与 A 的轨道间重叠的情况，如果我们取最简单的情况，即只有一个电子的体系，就很容易看到这一点，

$$\psi(\mathrm{D,A})=\psi_d;\ \psi(\mathrm{D^+—A^-})=\psi_a \tag{3.30}$$

式中 ψ_d 是 D 的给予电子轨道(即 HOMO)，ψ_a 是 A 的接受电子轨道(即 LUMO)。因此

$$\int\psi(\mathrm{D,A})\hat{H}\psi(\mathrm{D^+—A^-})\mathrm{d}\tau=\int\psi_d\hat{H}\psi_a\mathrm{d}\tau \tag{3.31}$$

从式 (3.31) 可以看出，除非轨道 ψ_d 与 ψ_a 彼此重叠，否则此积分为零；如果在某一点 ψ_d 或 ψ_a 为零，则在空间该区域 ψ_d 或 ψ_a 对此积分的贡献为零。Mulliken 指出，当 D 与 A 相互趋近形成“电荷迁移络合物”时，彼此间最可能的取向是要使 D 的 HOMO 与 A 的 LUMO 的重叠积分 S_{da} 有最大值。因为电荷迁移络合物(D,A)的基态能量 E_N 愈小，则(D·A)愈稳定，然而 S_{da} 愈大，E_N 就愈小，即(D·A)愈稳定。

在 Mulliken 提出电荷迁移络合物的价键理论后，便有人用分子轨道理论来描述给予体与接受体的相互作用[22]。其中，Flurry[23]提出的理论有可能广泛地用来计算实际体系。

Flurry 假定，D 与 A 作用生成电荷迁移络合物(D·A)的能量由两部分所贡献：一是形成分子间键时，D、A 的轨道相互作用能 E^b，作为一级近似只考虑 D 的 HOMO (ψ_d)与 A 的 LUMO(ψ_a)的相互作用；二是由于电子由 D 迁移至 A 而引起的静电相互作用能 E^e。

对于 E^b，在上述假定下，(D·A)的波函数 ψ 表为

$$\psi=a\psi_d+b\psi_a \tag{3.32}$$

在进一步假定 ψ_d、ψ_a 是正交归一的条件下，运用线性变分法解 (3.32) 式波函数 ψ 的 Schrödinger 方程，便可得到(D·A) 的基态和激发态的能量 E_N^b 和 E_E^b，以及相应的波函数 ψ_N 和ψ_E。在能量表达式中自然地引入了 D 的电离势 I_d，A 的电子亲合势 E_a 以及D、A的交换积分 β_{da} 等参数。

为了求得 E^e，Flurry 还假定，当电荷迁移络合物在平衡距离时，若一个电子完全从给予体 D 迁移到接受体 A，则其势能为 V。按照ψ_d、ψ_a 对总体波函数的贡献，(D·A) 在基态时，有 b^2 个电子从 D 迁移至 A，而在激发态时，有 a^2 个电子迁移。所以

$$E_N^e=-b^2V;\qquad E_E^e=-a^2V \tag{3.33}$$

Flurry 理论的能量关系图(见图 3-21)，其中量 $h\nu_{CT}$ 为电荷迁移跃迁能。

除了上述两种处理方法(Mulliken的价键法和Flurry的分子轨道法)外，将电荷迁移络合物作为一个整体，而运用第二章中所介绍的各种量子化学计算方法对(D·A)进行整体分子轨道处理也越来越受到重视。

我们知道，用实验方法测定生物学与药学上有意义的化合物的电离势和电子亲合势，是比较困难的。但是，HOMO和LUMO的能级却与电离势和电子亲合势有着十分密切的关系。对一系列给予体与同一接受体之间的相互作用，可用式(3.34)表示

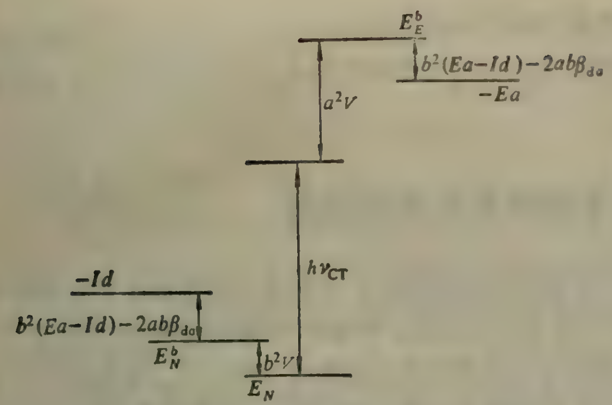


图 3-21 按Flurry理论电子迁移络合物中各能量之间的关系

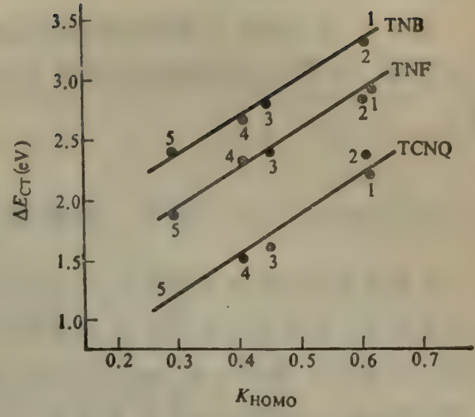


图 3-22 芳香烃类的电荷迁移络合物
电子受体:TNB(三硝基苯)TNF(三硝基氟苯)TCNQ*(四氰喹啉二甲烷)
电子给予体:1. 萘;2. 菲;3. 芘;4. 蒽;5. 四省。同理,对一系列受体与同一给予体,也有类似的关系。

$$\Delta E_{CT}=A-\alpha-k_{HOMO}\beta \tag{3.34}$$

式中,ΔE_{CT}是激发能,A是接受体的LUMO能量,α是库仑积分,β是共振积分。

对一系列烃给予体,用ΔE_{CT}与k_{HOMO}作图可得一直线(如图 3-22)。同理,对一系列受体与同一给予体,也有类似的关系。

3) 电荷迁移相互作用的生物学意义

如前所述,电荷迁移相互作用是生物系统中的重要作用方式之一。仅就络合物电荷迁移在激发能传递上的特点看,这些特点及其对DNA传递遗传信息的可能意义包括[24]:(1) 电荷迁移可分为自发的和光致的两种。前者又叫强电荷迁移,后者又叫弱电荷迁移。强电荷迁移不需要光能激发就可以自发地产生,因而并不需要很大的激活能就可形成激发能级[25]。自发电荷迁移络合物能发出相当于红外光量子能量的三线态辐射。因此,络合物电荷迁移传能的大小与生物体所能利用的能量子(例如ATP高能磷酸键的 10～12 kcal)大小十分一致,DNA 也就有可能利用这种传能方式来传递信息。(2) 络合物电子迁移使电子从一物质向另一物质迁移而能量并无多大损失,因为它并不引起分子结构或化学键的重新排列。这使我们联想到 DNA 复制和转录时, DNA模板分子链本身并不发生化学变化的事实(指作为模板时的一个单链)。利用电荷转移来传递信息,对不发生化学变化的模板是非常适合的。(3) 电荷迁移络合物是介于有规则闭壳分子和自由基之间的一种物质状态。络合物经电荷迁移后,如果离解成电子给予体和电子接受体的单分子,就可以形成两个具有非偶电子的自由基[20]。自由基活性很强,容易引起许多化学反应。

除以上所述外,电荷迁移相互作用还为半导体性进入生物学开辟了道路。同时也为了解生物氧化等生物学现象开辟了可能的途径。药学的研究还告诉我们,一些药物的作用是由于这些药物在某些关键步骤作为一个电子给予体或电子接受体,与生物接受体分子(或给予体分子)形

成电荷迁移络合物所致。

近年来，由于电荷迁移络合物理论的发展，使它已经成为研究分子间相互作用中不可缺少的理论基础。分子轨道理论的运用和量子化学计算的开展更推动了电荷迁移相互作用研究的发展。

3.3 生物分子的特异作用和识别

生物体内反应的最大特征之一，是具有高度的相互识别能力，这种能力的来源在于生物体内的分子间相互作用都是高度特异、高度协同和高度专一性的。例如DNA碱基间的配对(即A-T，G-C)、信息转录、翻译过程，密码与反密码的对应过程都存在高度的专一性，……。生物体识别非自身异物加以排除的免疫现象，也是基于抗原-抗体的专一性，而酶的专一性则表现在它们只对一些特定分子在特定位置上发生的特定反应起作用，这对于控制生物体内各种反应过程是极为重要的。

不言而喻，规定着这样的生物体内相互作用专一性的因子，无疑是非常复杂的，迄今为止还没有获得令人满意的解释。以下我们仅从量子生物学角度对这样的问题作一简述。

3.3.1 特异作用和生物分子识别的物理、化学基础

永田认为[3]，生物现象所见的高度专一性.生物体所进行的相互作用方式并不和非生物的截然不同，不过是在物理的、化学的水平上的专一性，以更复杂化、高度化了的形式表现而已。在这意义上，为要理解生物的专一性，有必要弄清楚非生物系统中相互作用的专一性。在此，我们不去评论永田的这些观点的正确性如何。但研究的实践表明，对物理的、化学的相互作用的专一性的研究，无疑有助于我们对生物专一性的了解。对生物分子(包括药物分子在内)大小、分子间力的作用和立体构型作用等的研究已经说明了这一点。其中，有关轨道对称性和能级的研究占据着一定的地位。

1）轨道对称性

Woodward-Hoffman发现了分子所具有的电子轨道的对称性，对于规定化学反应的专一性是很重要的。W-H法则的中心意思是“考虑反应前后如果保持了轨道对称性，对称性容许(symmetry-allowed)则进行反应，对称性禁止(symmetry-forbidden)则不起反应”。亦即，反应物的分子轨道必须以对称性守恒的方式转化为产物的分子轨道。当反应物的成键轨道与产物的成键轨道相关，基态分子的直接变换是可能的，这个途径的活化能低，被称为对称性容许的反应。如果反应物的成键轨道与产物的反键轨道相关，轨道对称性抑制分子变换，即反应要遇到一个反映激发态能量的势垒，这个途径活化能高，被称为对称性禁止的反应。

W-H法则不仅对分子内反应，对分子间反应也广泛适用，并取得了很大的成果。有关W-H法则较详尽的介绍可参见文献[3,26]。现在以最简单的例子：两个乙烯分子生成环丁烷的加成反应来说明轨道对称性守恒原理的处理方法。

二个乙烯分子生成环丁烷的反应，由HMO法得乙烯的分子轨道为：

$$\psi_1=\frac{1}{\sqrt{2}}(\phi_1+\phi_2),\quad \psi_2=\frac{1}{\sqrt{2}}(\phi_1-\phi_2)$$

ψ_1为成键轨道，轨道上有二个电子；ψ_2为反键轨道（空轨道）。当二个分子接近时，具有相同对称性的 $\psi_2^{(1)}$与 $\psi_2^{(2)}$能形成结合状态的轨道 $\psi_2^{(1)}+\psi_2^{(2)}$与 $\psi_2^{(1)}-\psi_2^{(2)}$，同理 $\psi_1^{(1)}$与 $\psi_1^{(2)}$也能形成结合状态的轨道$\psi_1^{(1)}+\psi_1^{(2)}$与$\psi_1^{(1)}-\psi_1^{(2)}$。

假设反应进行时二个乙烯分子以相互平行的分子面形状靠近，则显然把 它们整体 视作一个大分子时，这个准分子具有C_{2v}群的对称性。这个简单的对称群具有E、C_2、σ_v、σ_v'四个元素，E为全等变换，C_2、σ_v、σ_v'分别对应图 3-23 中的C_2，m_1和m_2。

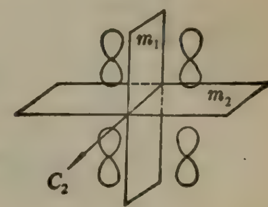


图 3-23　二个乙烯分子反应时的对称性示意图

对照C_{2v}群表示，我们可以得出二个乙烯分子反应时轨道波函数能级分布及其轨道对称性。而新形成的产物——环丁烷的轨道是二个反应分子的轨道变形融合而成，即二个轨道 适当取向重叠而成。根据对称性守恒，要求新的分子轨道保持原有轨道的对称性。 作 出轨道相关图如图3-24所示。

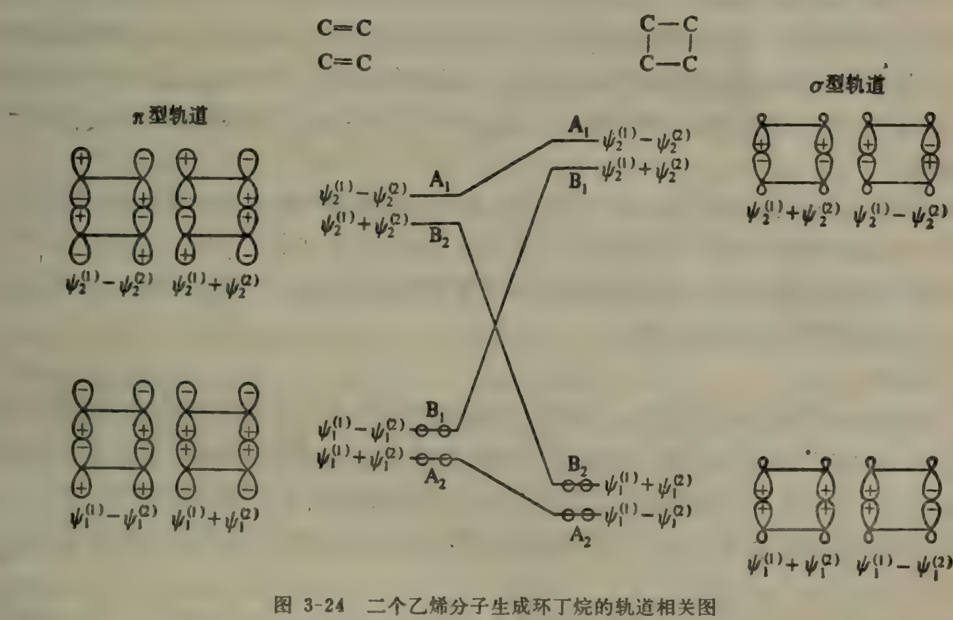


图 3-24　二个乙烯分子生成环丁烷的轨道相关图

图中告诉我们，反应对基态是对称性禁止的，只有在激发态才满足对称性容许，说明 反应需较大的活化能才能进行，加热不能进行反应，实验结果也确实是如此。

2）能级[3]

研究表明，电子的性质与分子间相互作用的专一性密切相关。我们可以举 长仑·田仲 提出的用电子能级解释化学反应为例。

长仑·田仲比较亲电子、亲核、自由基的各种试剂和被反应分子的电子能级，发现了一些

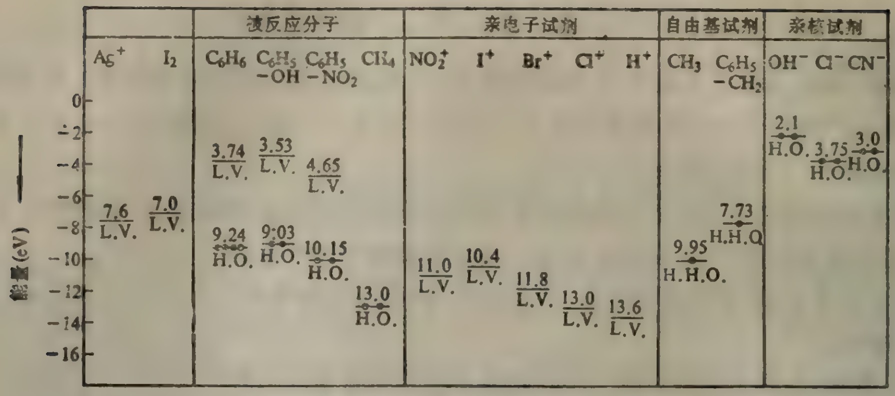


图 3-25 被反应分子和各种反应试剂的能级图
H.O. 表示最高占据能级，H.H.O. 为半占据能级，L.V. 为最低空能级

规则(见图3-25)：

(1) 亲电子试剂的最低空能级低于被反应分子的最高被占能级时发生反应（如 NO_2^+，I^+，Br^+，Cl^+等对苯，酚，硝基苯的反应），高于被反应分子的最高被占能级时则不发生反应（如CH_4和NO_2^+，I^+等的反应）；(2) 亲核试剂的最高被占能级高于或接近被反应分子的最低空能级时，发生反应；(3) 反应发生时，自由基试剂的最高半被占能级在被反应分子的最高被占和最低空能级的中间；(4) 如Ag^+和I_2那样，最低空能级高于被反应分子的最高占据能级时，不发生反应，而生成分子化合物。

这些规则对于说明生物体内发生的特异相互作用和识别究竟能达到何等程度，还难以作出结论。不过，可以预料分子间相互作用的专一性，在一定程度上与电子能级之间是可能存在着某种相关关系的。我们可以通过分子轨道法的计算得到参加反应的分子的电子能级，进而探讨电子能级对分子间相互作用专一性的影响。

3) 空间构象

分子生物学的研究成果告诉我们，生物分子间的特异相互作用和识别与分子(包括药物分子)本身的空间构象有着相当直接的关系(参见后续有关章节)。例如药物分子本身的构象要与受体的构象相适应，才易于与受体结合而发挥作用。遗传信息的表达所受到的严格调控，是同DNA与其它分子的相互识别、相互作用直接相关的，是通过DNA的微妙构象变化实现的等等。

量子化学可以通过计算分子的能量来推测分子的构象，因为在不同的构象中，各个相互作用的原子或基团间距不一，从而导致整个分子能量的不同。能量较低的构象比能量高的构象更为稳定，高能量的构象可能自动转化为低能量的构象。同样，在药理学的研究中业已知道，与药物分子相作用的生物体大分子，也稳定在低能量的构象，但大分子化合物的结构更为复杂，有时几个不同构象的能量都较低，并且互变的能障也不大，几种构象就可并存，大分子也易于变形。药物或内源活性物质与这大分子的不同构象作用，分别产生不同的效应。

有关空间构象的分子轨道研究的讨论，读者可参考第四、第八和第九章中的有关部分。

3.3.2 酶与底物的专一作用

我们知道,酶依靠有组织的,错综的空间结构,使得其表面具有一些特殊区域。这些区域依据它们的三维形状,电荷特征及疏水与亲水程度的不同,而使它们能与特定的底物相结合。底物一般是通过非共价作用与酶的这个特定区域相结合而成为酶-底物的复合物。这个特定区域的附近或内部还同时存在一定性质的催化基团,这些催化基团被巧妙地安排在酶侧链的一些关键位置。

酶反应专一性的概念包括(1) 反应专一性,即只催化一类反应;(2) 底物专一性,即只催化一种或几种底物的转化;(3) 束缚专一性,只束缚一种底物到酶的活性部位上;(4) 动力学专一性,即底物束缚到酶活性部位之后,在催化阶段底物反应性上的差别。

以下我们仅就酶对底物专一性的基础究竟是什么这样的问题作一初步讨论。

3.3.2.1 活性中心与诱导-契合学说

已经知道,参与酶活性中心的化学基团实际上就是酶蛋白中氨基酸残基的侧链或肽链的末端氨基与羧基。这些基团一般不集中在肽链的某一区域,不一定互相邻近,有的甚至可分散在不同的肽链上。例如胰凝乳蛋白酶活性中心的五个氨基酸残基为异亮氨酸 16,组氨酸 57,天冬氨酸 102,天冬氨酸 194 与丝氨酸 195,它们分别分散在二条肽链上,主要是依靠酶的二级、三级结构才使它们集中在活性部位所在区域,正是这个区域导致了胰凝乳蛋白酶的专一性。

我们知道,在胰凝乳蛋白酶的全部氨基酸残基中共有 27 个丝氨酸,但实验发现只有一个丝氨酸,即 195 号丝氨酸存在于活性中心。为什么它会成为活性中心呢?通过量子生物学研究了解到,原来是它所处的环境,使得周围的负电荷向它移动,使其电负性增加而成为活性的亲核试剂。

1958 年,Koshland 提出了"诱导-契合"假说[27]。他认为活性中心是挠性的而不是"锁与钥匙"理论中认为酶分子的结构是刚性的。这个假说提出:在底物达到酶分子结合部位之前,并不存在一个精确互补的结合部位,而仅仅是一个相结合的倾向而已。但一旦由于这种倾向而吸附了底物之后,这时底物周围的结合部位才越来越牢固地束缚底物并限制了它的自由度,直至二者之间的契合十分贴切。对这二者的契合过程一般有两种解释:一种是酶分子形变,结合基团的移动以达到协调。另一种是认为与不精确互补的活性部位结合的部分内能被用来使底物形变,酶将会利用潜在的结合能将底物推向过渡态结构。换言之,当酶结合底物后就会在底物结构上诱导应变以导向过渡态结构,这样就促进了反应的进行。这些应变力的确切本质尚不能很好地下定义,看来它们涉及酶与底物间的非共价键的相互作用。这种应变理论被归为"现代架子假说"之内。有时也有二种情况同时发生的,如溶菌酶即是。

溶菌酶尤如一只鸟,两只翅膀之间的裂缝即为其活性部位,如图 3-26。

羧肽酶 A 与甘氨酰酪氨酸的结合情况更可说明诱导-契合在动态地识别底物方面的作用。这时精氨酸 145 带正电荷的胍基移动 0.2nm,谷氨酸 27 的羧基也有类似的移位,而最大的构象改变为酪氨酸 248 的羧基移动了 1.2nm,从亲水的分子表面移位到底物肽链附近的疏水区域。这种种的构象改变都是由精氨酸 145 结合到底物末端羧基时开始发生的。如果底物中没有游离

的末端羧基，则酪氨酸 248 不能移位，底物不能贴切嵌入，催化作用也无法产生。这说明诱导-契合在动态地识别底物，决定底物专一性方面具有重要的作用。

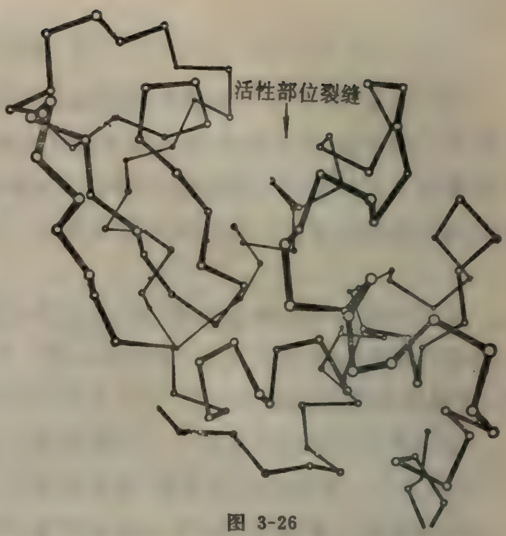


图 3-26

无数事实说明了酶的活性部位由于基团特定结构及各基团的巧妙配合，通过诱导-契合作用而构成了酶分子的高度专一性，活性部位中接触残基与结合基团的作用决定了底物专一性。又由于酶至少三点接触底物才能辨认，由此产生了立体专一性。

3.3.2.2 酶专一性本质的进一步探讨

众所周知，酶对外界反应极灵敏。外界条件的细微改变，都可能造成酶侧链间力的平衡遭到破坏，引起其构象的变化，进而导致活性的变化。为进一步弄清活性部位的作用，有必要对形成酶构象以及引起构象变化的各类作用因素加以研究，进而挖掘出专一性现象的部分微观本质。

1）疏水性相互作用

由于水的结构内没有容纳非极性残基的能力，使酶的非水溶性侧链具有避水倾向而聚拢到折迭肽链的"核心"，于是使亲水性侧链留于表面而适宜与水性环境相互作用，疏水基团与疏水侧链自动靠拢造成"疏水键"。该键对酶的专一性有一定作用，例如 α-糜蛋白酶作用于底物的氨基酸部分的专一性，就取决于糜蛋白酶表面非极性部分与底物部分所形成的"疏水键"。

可以说，熵是造成疏水作用的动力，酶的疏水侧链正是通过形成非极性的受水包围的小团来防止有序性的增加和熵的减少。显而易见，在保证水的有序结构的温度范围内，温度越高，疏水作用将越强。因为由$\Delta E=\Delta H-T\Delta S$知，随着 T 增加，$T\Delta S$ 项变得愈来愈不利于非极性基团的溶解。

Knowles 由水解酶的动力学数据得出结论，认为高疏水性，加上电荷条件的满足，会增加底物的束缚和转化的专一性。根据这种分析，胰凝乳蛋白酶的活性中心应是非极性的。乙酰胆碱酯酶和胰酶等专门劈裂带正电的分子，对具有大的疏水残基的底物增加催化作用。

2）静电相互作用

静电力首先可以形成氢键，在肽的C═O 基与各侧链基团之间，其中特别是一个肽的 C═O 基与另一个肽的 N—H 基之间，及侧链基团本身之间都可形成氢键。如果说酶的结构以允许有最大的疏水方式进行折迭，则可以说，与此同时，各种侧链及肽骨架间形成的氢键与该作用完全一致。另外，酶侧链的带电基团之间也可能产生静电相互作用，例如 N末端残基或赖氨酰侧链的 NH_3^+，或精氨酰侧链上的胍基，组氨酰残基的咪唑基上的一个正电荷都可被谷氨酰或天冬氨酰残基侧链上羧基—COO^-的一个负电荷以静电力所吸引。如果这些带电基团都伸入酶周围的水性介质中去，便可在其周围形成稳定的水合层。当在酶溶液中加酸而降低 pH时，硫醇盐—S^-和羧酸盐基—COO^-将逐渐质子化，变为不带电的硫氢基—SH和羧基—COOH，而氨基、咪唑基

和胍基将保持带正电。反之,加碱前者将保持其负电,而氨基与咪唑基将逐渐失去质子。这些不同的电荷特征都可导致酶构象的变化或影响其结合底物的特异性，这些不同的电荷特征也从一方面决定了酶的专一性。

3) Van der Waals 力

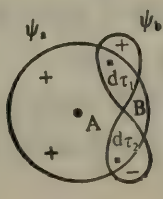
一对原子间的Van der Waals 力是较小的,但当底物的许多原子同时接近酶的原子时，Van der Waals 力的作用不可忽略。换言之，底物与酶之间有效的Van der Waals 力只有当它们处于立体互补时才能产生,这就说明虽然在某单一的Van der Waals 力作用中，并不存在专一性，但当同时形成大量的Van der Waals 相互作用时就必然出现专一性。

酶与底物的可逆结合是通过三种不同类型的最基本的键作用来实现的,这三种键即离子键，氢键和Van der Waals 键,它们可以由于水的存在而受到各种影响。同时，它们对几何形状的要求，其强度及专一性方面有所不同。正是上述的这些相互作用关系各自起作用或者重叠起作用,使酶能与底物以最适当的键结合,从而把底物稳定地固定在特定结构的结合部位上，体现出酶分子的高度专一性。

3.3.2.3 从轨道对称性看酶的专一性

由量子力学的基本观点可知，交换积分 H_{ab} 的值在原子结合成分子中起着关键的作用。而 H_{ab} 与重叠积分 S_{ab} 有密切关系,具体地说,要有效地组成分子轨道，必须要求 S_{ab} 越大越好,还要求必须以一定的取向——满足对称性要求的前提下相互结合才行。

例如当一个 s 轨道(ψ_a)与 p 轨道(ψ_b)作用时，即使像图 3-27 (A) 那样双方靠得近,电子云重叠部分很大,但二者无法成键。这里因为对任一微体积元 $d\tau_1$，必然可以找到其对应的体积元 $d\tau_2$,二者对 S_{ab} 的贡献正好抵消,因而 S_{ab} 为零。 但如图 3-27 (B) 所示的二个轨道，则可以结合。化学反应在轨道对称性方面的要求对专一性问题应该说也有一定的影响。



A

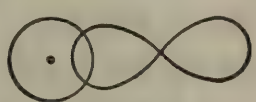

B

图 3-27 s 轨道和 p 轨道的重叠方式

前面已提及的 W-H 法则由于注意到了电子轨道的对称性，因而对于规定化学反应的专一性是极为重要的。 Ferreira 在应用轨道对称性的设想来解释酶反应机制时,也提出了饶有兴趣的想法[28]。

3.3.2.4 酶与辅酶间的相互作用[3]

我们知道,全酶是酶蛋白和辅酶的复合物。 酶离开辅酶是不能完成对底物的催化反应的。那么,酶与辅酶间究竟通过怎样的相互作用才能实现酶的催化功能呢？在这方面研究得最多的辅酶是辅酶 I(NAD)和辅酶 II(NADP)。

从图 3-28 可以看到，NAD 和 NADP 之间在结构上只差一个分子磷酸，故两者的作用类似，均系脱氢酶的辅酶。

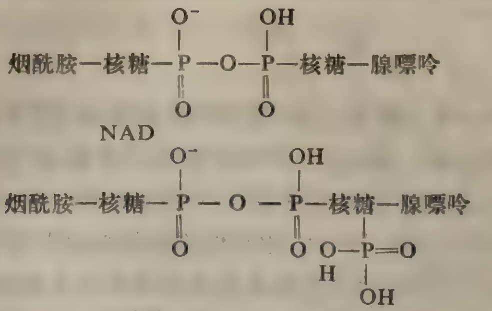


图 3-28　NAD 和NADP 分子示意图

当 NAD 和酶蛋白混合时，在 360 nm 附近出现极大吸收(一条宽带叫 Racker 带)。这条带在了解辅酶和酶蛋白相互作用形式方面非常重要。已经知道，烟酰胺部分是脱氢作用时重要的氢的交接场所(见图 3-29)。由图可见，在吡啶核的 4 位，底物的 H 以阴离子形式与之加成，起到了脱氢作用。所以 4 位对亲核基团的反应性要求大。

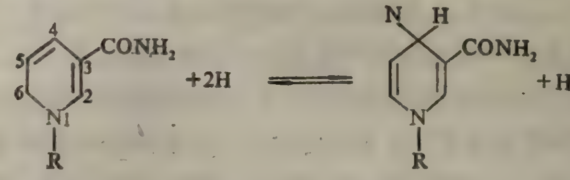


图 3-29　NAD 的氧化型(左)和还原型(右)

用 P-P-P 法计算的 NAD^+烟酰胺部分全 π 电子密度以及在亲核反应时前沿轨道（最低空轨道)的电子密度，如图 3-30 所示(图中“·”为反应性最大位置)。

由图 3-30 可见，全 π 电子密度和前沿电子密度的值很好地反映了上述情况，如 4 位的全 π 电子密度值最小，另一方面，最低空轨道的电子分布，即在亲核基团前沿电子密度是最大的*。

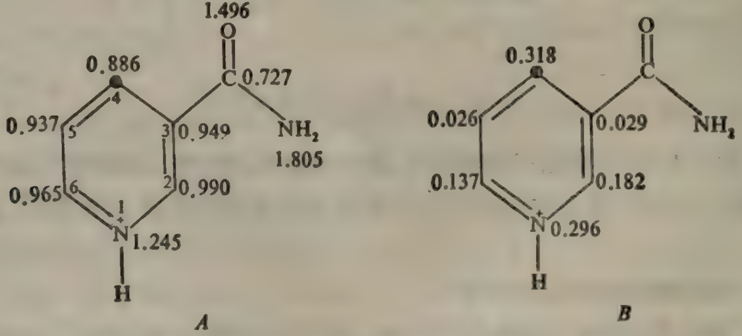


图 3-30　NAD^+烟酰胺部分全 π 电子密度(*A*)和亲核反应时前沿轨道(最低空轨道)的电子密度(*B*)

* 对于亲核试剂和反应性的关系：全 π 电子密度小，最低空轨道的电子分布比较密，活性大。

因此要很好注意到，电子结构论认为最具活性的位置正是酶反应的活性位置。实际上，许多对烟酰胺反应的亲核试剂全部在4位上起加成反应。

如上所述，当NAD和酶蛋白混合时，在360 nm附近出现极大吸收的宽吸收带。迄今为止，对于这一吸收带的起因仍旧存在着不同的看法。其一是Racker等[29]主张的，他们认为是酶蛋白的SH基和烟酰胺的吡啶环4位起加成反应的光谱。其根据是NAD和各种亲核试剂(OH^-, CN^-, SO^-_2, $NHOH^-$等)容易在吡啶环起加成反应，而且加成物的光谱在360 nm有吸收。对此，Kosower从吸收带宽而弱看来，认为和加成物相比，考虑酶蛋白和辅酶之间的电荷转移更为合理[30]。

Umeyama等[31]则认为，NAD形成电荷迁移复合物是酶蛋白的色氨酸。那是因为当NAD和色氨酸混合时，在320 nm附近出现新的吸收带，可以认为这是由于二者相互作用的结果，而其他氨基酸则完全看不到这种相互作用。与酶蛋白的色氨酸残基相互作用的NAD的作用基团是烟酰胺的吡啶环，从引起的光谱变化看来是很清楚的，但与色氨酸的吲哚核以怎样的配置方式相互作用则还不很明了。

不过为了产生电荷转移，考虑吡啶环和吲哚环以堆积形式相互作用最为妥当。在各种堆积形式中，最稳定的结构而且它的最大吸收和实验值很好符合的配置如图3-31。实线表示烟酰胺部位，虚线表示吲哚环。



图 3-31 吲哚环和烟酰胺部分的堆积

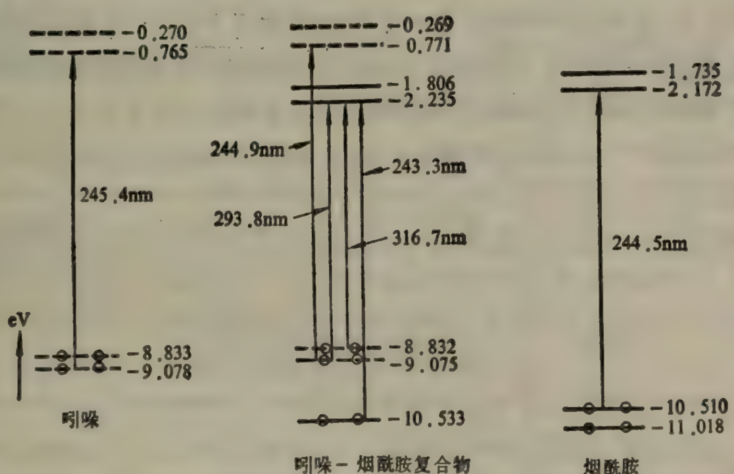


图 3-32 吲哚、烟酰胺和它们的复合物的能量能级和跃迁烟酰胺以N_1位置上加成H的结构代替——烟酰胺，……吲哚

按照这种配置，用P-P-P法计算能级和跃迁的大小，以及它们和无堆积时吲哚和烟酰胺的比较如图3-32所示。复合物的电荷迁移带在316.7 nm（注意：在接近293.6 nm也有），和实验值320 nm大体上一致。因而确认了NAD和色氨酸相互作用出现在320 nm的吸收带是建立在后者的吲哚环与前者的烟酰胺电荷转移这一理论上的。这也强有力地支持了Kosower

的设想，即 NAD 和酶蛋白相互作用产生的 360 nm 附近的光吸收是基于从酶的色氨酸残基向 NAD 的烟酰胺部分的电荷转移。但是，NAD-酶蛋白的吸收带是在 360 nm，比 NAD-色氨酸的情况在长波差 40 nm 波长。这种差别可能是由于酶存在电离的侧链，因而存在电场，有可能引起光谱位移。当然，在这种情况下，电场引起移位只是许多因素中的一个，也可能还有其它因素引起移位的可能性。

以上我们介绍了永田等运用分子轨道法对 NAD 的研究。鉴于有关酶与辅酶间的相互作用对于体现酶的特异相互作用和识别，以及高效催化率具有十分重要的作用，因而近些年来，在实验和理论研究两个方面都取得了较大的进展。通过上述以辅酶中 NAD 为例子的简述，我们可以看出，分子轨道法在分析酶和辅酶间的相互作用中，提供了某些与实验结果很好符合的理论计算结果，可以预期分子轨道法在这个领域将获得更广泛的应用。

3.3.3 密码子与反密码子的识别和相互作用

我们知道，在蛋白质生物合成过程中，密码子和反密码子的相互识别是非常精确的（误差约为 10^{-3}—10^{-4}）。但是这种高度专一的相互识别的机制究竟是什么呢？它与密码子和反密码子配对缔合时碱基三联体间相互作用能的关系究竟如何？怎样从分子间相互作用力理论来阐明碱基异构体参与配对的可能性？等等，这些都是我们所关心的问题。

近几年来，一些研究者从实验和理论两个方面对这些问题进行了探讨。例如，我国科学工作者江寿平等[15]根据经验的多原子分子间相互作用势研究了氨基酸的遗传密码子与反密码子的相互识别。他们以RNA12 的几何构型作为理论模型，计算了RNA12 上互补碱基间，非互补碱基间和碱基三联体间的相互作用能，并对碱基异构体参与的非互补碱基配对的相互作用能函数进行了自选尺度拟牛顿法的最优化处理，结果表明(见表3-5，表3-6)。

表 3-5　RNA12中互补碱基共面配对的相互作用能（kcal/mol）

	总　　能	静　电　能	极　化　能	色散能＋排斥能＋氢键能
A·U (U·A)	−7.17	−5.07	−1.70	−0.40
G·C (C·G)	−26.55	−20.92	−6.21	0.58

(1) RNA12上互补碱基共面配对相互作用能的贡献主要来自于它们的静电相互作用(表 3-5)。而非互补碱基间配对的相互作用能计算均大于互补碱基间配对的相互作用能至少6.49 kcal/mol(表 3-6)($\Delta E = E_{非互补} - E_{互补}$，这是假、真反密码子同密码子相互识别能力的强弱度量因子，其界限约为 5.6 kcal/mol)。因此，根据分子识别与其相互作用能的关系，可知在RNA12中位于不同螺旋链的碱基，其相互识别配对是严格遵循标准的碱基配对原则，即碱基A·U和G·C间的相互识别而配对。这种识别配对的本质在于互补碱基A·U，G·C间高度的专一相互作用，在于它们的配对具有最负的相互作用能。这在一定程度上也表明，在生物分子的识别过程中，生物分子间的直接的物理相互作用起着十分重要的作用。

表 3-6 RNA12中非互补碱基共面配对相互作用能（kcal/mol）

	最小原子间距（Å）	糖苷间距（Å）	相互作用总能	静电能	极化能	色散能＋排斥能
A·C（C·A）	1.013	10.496	29.56	11.27	−4.95	23.24
A·G（G·A）	0.191	10.518	9.86×10^{12}	−162.21	−17428.27	9.86×10^{13}
A·A	0.420	10.516	1.43×10^{8}	35.96	−357.38	1.43×10^{8}
U·C（C·U）	3.105	10.476	−0.68	0.39	−0.33	−0.75
U·G G·U	0.985	10.498	5152.60	6.31	−6.88	5153.17
U·U	2.741	10.476	2.34	3.12	−0.08	−0.70
C·C	2.166	10.476	7.04	8.41	−0.48	−0.89
G·G	0.654	10.520	1.76×10^{13}	4.86	−4921.36	1.76×10^{12}
G·I（I·G）	0.676	10.411	2.53×10^{13}	−49.57	−6356.60	2.53×10^{13}
U·I（I·U）	0.991	10.400	4821.37	9.49	−4.54	4816.42
C·I（I·C）	1.903	10.400	−15.67	−11.69	−3.98	0.01
A·I（A·I）	0.201	10.409	6.28×10^{12}	−147.88	−14902.77	6.28×10^{12}

（2）在有碱基异构体参与的非互补碱基间配对的情况下(参见图 3-33 和表 3-6，3-7)，通过对相互作用能的计算及优化处理，结果表明，当允许碱基配对存在一定自由度时，则存在着某种异构体参与配对的可能性。

关于碱基异构体存在时的配对方式已在 4.2.2 中讨论过(参见7.2)，Полтев等[32]在用原子-原子势函数对包含两个碱基和一个水分子的系统作分子间相互作用能的计算时，列出了29种通过水分子而建立的碱基配对方式（图 3-33）。由于异构体具有和正常型不同的氢键专一性，所以产生正常型 Watson-Crick 型以外的氢键对存在的几率，如鸟嘌呤，肌苷的异构体和尿嘧啶配成对，尿嘧啶的异构体能和鸟嘌呤、尿嘧啶配成对，其结果得到如表 3-7 所示的配对，Crick从Wobble学说推导出和密码子-反密码子的碱基对的多样性完全相同的结果。

以上所述再次提示我们，对所谓“DNA复制和转录时选择对方的基础，早已清楚是在于 A—T(U)，G—C这些氢键的专一性。”的见解似乎应该持保留态度。

我们知道，DNA分子双螺旋结构稳定性是复杂的，远不能简单地认为是氢键所致。作者认为，所谓“氢键的专一性”的提法是含糊的、不够严格的。一种说法是：正常型碱基配对系依赖于所谓的氢键专一性，而异常型的存在则影响了碱基对的氢键专一性。这里既然没有指出氢键专一性的实质，那末正常(或异常)碱基的配对，就难以用“氢键专一性”来判断。

鉴于迄今为止，我们对碱基配对的氢键专一性的实质尚不甚了解这一实际，因而以电子结构理论理解遗传的密码子与反密码子的识别和相互作用的尝试，就具有一定的意义。

1966年Crick根据立体化学原理，提出了Wobble学说(参见7.1.2)，认为反密码子在识别密码子的过程中，密码子与反密码子之间的前二对碱基严格遵循标准的碱基配对原则，而它们的第三对碱基则允许存在一定的自由度。在互补、非互补碱基配对相互作用能计算的基础上，江寿平等

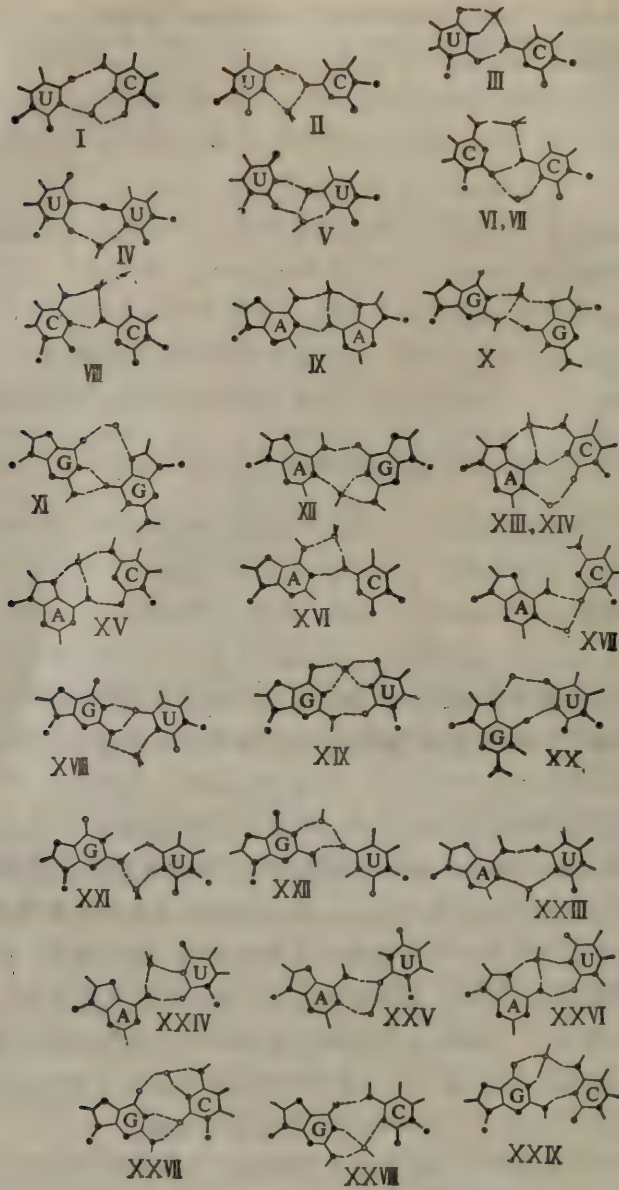


图 3-33

进一步计算了碱基三联体(密码子与反密码子)间的相互作用(见 96 页表 3-8),结果概括为:

(1) 密码子的第三个碱基U能与反密码子中碱基A,G(醇)及I(次黄嘌呤,醇式)配对，但不能与G、I的酮式正常型配对。

(2) 密码子的第三个碱基A(顺式)能与反密码子中的碱基I 配对。

(3) 碱基A·U和G·C配对可以在密码子,反密码子中的任何位置进行。

这就在一定程度上反应了 Crick 的Wobble 学说的机制,并从相互作用能的角度说明了现有的大多数密码子和反密码子间识别关系的实验资料。

表 3-7 考虑了异构体结构时的Watson-Crick型氢键对的反密码子和密码子的碱基对

反密码子上的第一个碱基	密码子上的第三个碱基
U	A
U*	{G, U}
C	G
A	U
G*	U
G	C
I*	U
I	{C, A}

* 符号表示异构体，I 表示次黄嘌呤

3.3.4 抗原-抗体反应的专一性

生物体具有识别非自身异物加以排除的免疫能力。当这种外来异物作为抗原侵入生物体后，生物体的免疫机构就会产生抗体与抗原形成沉淀反应，达到免疫效果。

每种免疫细胞在其表面都具有抗体结合部位，起到识别的作用，并探查外来的抗原决定因子，这种情况即使在外来抗原物质浓度很低时也会发生。然而，它们并不能识别自身抗原决定因子，有时甚至浓度很高时，也是如此。抗原决定因子是由不太大的寡聚物组成的，上述情况的发生很可能是由于自身抗原决定因子的结构特点（蛋白质的内部排列）及其他因素造成的。

抗原是非自身的多肽和多糖类物质。通常它的分子量最低界限为 4 000～10 000，偶尔也有1000左右的抗原物质。抗体是球蛋白类。抗体结合部位约有10～20个氨基酸，大小与抗原决定因子差不多。

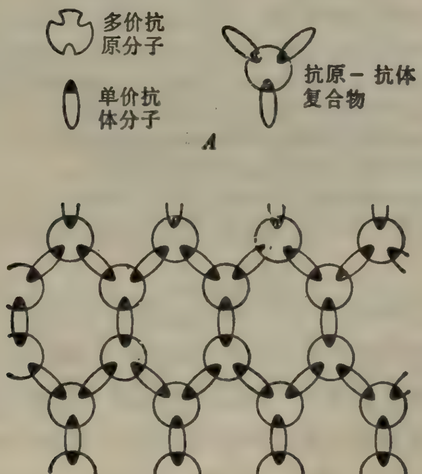


图 3-34 单价抗体与多价抗原之间的反应（*A*），二价抗体与多价抗原之间的反应（*B*）

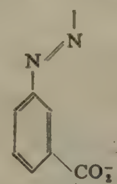
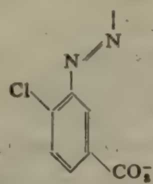
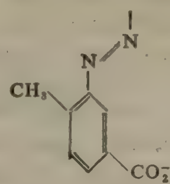
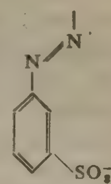
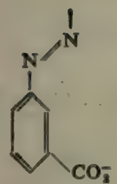
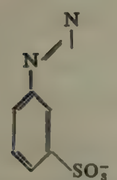
表 3-8 部分的密码子与反密码子相互作用能(kcal/mol)

密码子 5′→3′	反密码子(真，假)3′→5′					密码子 5′→3′	反密码子(真，假)3′→5′				
AUG	UAC	UCC	CUC	CCC	UAU, UAA	GAC	CUG	CUI	CUU	CUC	CUA
	−43.70	−35.29	−8.74	−0.32	强排斥		−60.06	−46.97	−30.68	−20.17	−0.96
GGG	CCC	CCU*	CCU CCG CCA CAC			CGC	GCG	GCI*3	GCU	GCC	GCA
	−94.48	−68.22	强排斥				−68.17	−57.68	−40.39	−38.97	−16.33
GGA	CCU	CCC	CCI CCG CCA CUU			CGA	GCI*4	GCU	GCC	GCA GCG	
	−62.91	−32.86	强排斥				−48.83	−47.47	−17.04	强排斥	
GGC	CCG	CCA	CCC	CCU	ACGCAG	CGU	GCA	GCG*1	GCI*2	GCC	GCU
	−87.92	−32.22	−53.76	−55.79	强排斥		−53.93	−50.71	−45.96	−47.56	−38.24
GGU	CCA	CCG*1	CCU	CCC	CAA	UAA	AUU	AUC	AUG AUA AUI		
	−71.55	−71.52	−54.38	−64.07	强排斥		−20.43	18.43	强排斥		
GCU	CGA	CGI*2	CGC	CGU	CGG	UAG	AUC	AUU*	AUA AUG AUI		
	−57.86	−54.99	−46.12	−48.65	强排斥		−37.39	−19.95	强排斥		
GCC	CGG	CGI*3	CGA	CGU	CGC	UGG	ACC	ACU*	AUA AUG AUI		
	−85.92	−72.68	−24.49	−55.03	−41.76		−71.14	−44.22	强排斥		
GCA	GCI*4	CGU	CGC	CGA CGG		UGU	ACA	ACG*1	ACC	ACI*2	ACU
	−63.39	−61.27	−18.99	强排斥			−47.67	−45.65	−41.27	−39.23	−30.91
UCA	AGI*4	AGU	AGC	AGA AGG		UGC	ACG	ACI*3	ACU	ACC	ACA
	−41.07	−39.64	1.98	强排斥			−61.27	−48.46	−31.83	−31.45	−8.84
UCU	AGI*2	AGA	AGU	AGC	AGG*1	GUU	CAA	CAG*1	CAI*2	CAC	CAU
	−32.42	−35.61	−26.82	−24.95	−29.94		−46.53	−39.49	−38.83	−37.55	−32.06
UCC	AGG	AGI*3	AGU	AGC	AGA	GUG	CAC	CCC	CUC	CAU, CAG	
	−62.91	−50.58	−33.70	−21.09	−2.75		−66.98	−60.89	−47.42	强排斥	
AUU	UAA	UAI*2	UAC	UAG*1	UAU	GUC	CAG	CAI*3	CAC	CAU	CAA
	−23.43	−16.06	−15.22	−14.85	−10.14		−60.06	−48.12	−28.13	−33.38	−8.10
AUC	UAG	UAI*3	UAU	UAC	UAA	GUA	CAI*4	CAU	CAC	CAA CAG	
	−37.65	−26.54	−12.24	−6.58	14.22		−41.97	−37.18	−2.92	强排斥	
AUA	UAI*4	UAU	UAC	UAG UAA		GAG	CUC	CUU CUI CUG CUA			
	−19.84	−19.59	15.10	强排斥			−54.93	强排斥			
GAU	CUA	CUG*1	CUI*2	CUC	CUU						
	−37.65	−34.52	−33.25	−27.85	−27.62						

* 反密码子中碱基U与密码子中第三个碱基为鸟嘌呤醇式异构体的配对。
*1 反密码子中碱基(5′端)为鸟嘌呤的醇式异构体且处于最优配位位置。
*2 反密码子中碱基(5′端)为次黄嘌呤的醇式异构体且处于最优配位位置。
*3 反密码子中碱基(5′端)为次黄嘌呤的酮式正常型且处于最优配位位置。
*4 反密码子中碱基(5′端)为次黄嘌呤的酮型且与密码子中第三个碱基A(顺式构象)配对并处于最优配位位置。

有证据指出[33]，抗原分子就象蛋白质一样是多价的(亦即它具有许多与相应抗体结合的结合部位)。由于大多数抗原-抗体复合物是不溶解的，这似乎意味着，抗体之所以发挥它的功能，是由于与抗原分子交联，产生出能从溶液中沉淀出来的无限大的分子尺寸的复合物。如果抗体是单价的，在这种情况下虽然也能形成小的、可溶性的抗原抗体复合物[图 3-34(*A*)]，但是不可能发生交联。要使抗体与抗原分子发生交联，抗体结合部位的最小数值是二。如果我们设想类棒状的抗体分子正好在每一端上各有一个结合部位，则交联是可能发生的[图3-34(*B*)]。

表 3-9 抗原-抗体反应中半抗原的结构变化的效果

制备抗体时所用的马血清抗原上的半抗原	鸟血清抗原上的半抗原					
						
	0	+++	++++	+++	0	+
	0	0		0	0	++++

\+ 表示产生沉降反应，0 表示不产生。

鉴于对抗原-抗体反应专一性的研究涉及到大分子问题，因而给量子力学计算带来了很大的困难。但是，我们已经知道，生物大分子间、生物大分子和其它分子间的相互作用是明显地伴随着构象变化的。在抗原-抗体反应中，专一性的抗体既可以引起作为抗原的蛋白质构象的变化，也可以稳定抗原的特殊构象。由此可见，如果能从“构象”的处理入手，我们就有可能获得一些有意义的结果。再者，从表 3-9 可以看到，抗原-抗体反应的专一性，很大程度上依赖于抗原上连结的半抗原的性质。即，芳香环上间位有CO_2^-或SO_3^-基等半抗原的马血清抗原所制作的免抗血清(抗体)和鸟血清抗原上半抗原反应，则鸟血清抗原上半抗原，同样也必需在间位上有 CO^-_2，SO_3^-基，这些基团在对位，或者不具有负电荷基团时，则全然不发生反应。这些事实表明：规定抗原-抗体反应的专一性的因子，是和抗原、抗体分子的某特定部位的立体条件和电子性质有密切的关系。

参考文献

[1] Давыдов, А. С.: “Биология и Квантовая Механика”, (Издательство «Наукова думка», Киев), (1979).

[2] L. 鲍林著，卢嘉锡等译：《化学键的本质》，上海科学技术出版社，(1966).

[3] 永田亲义著，陶宗晋、江寿平译：《量子生物学入门》，上海科学技术出版社，(1979).

[4] 徐光宪：《物质结构》，人民教育出版社，(1961).

[5] Claverie, P.: “Intermolecular Interaction from Diatomics to Biopolymers”, Pullman, ed (Wiley, New York), (1978).

[6] Rein, R. et al.: J. Theor. Biol., 34:215 (1972).

[7] Amos, A. T. et al.: Theor. Chim. Acta, 40:221 (1975).

[8] Langlet, J. et al.: Inter. J. Quantum Chem., 20:223 (1981).
[9] Hirschfelder, J. O. et al.: "Molecular Theory of Gases and Liquids". (Wiley, New York),(1954)
[10] Kang, Y. K. et al.: Theor. Chim. Acta, 61:41 (1982).
[11] Ahmed, N. A. et al.: Acta Cryst., B28:739 (1972).
[12] Caillt, J. et al.: Acta Cryst., A31:448 (1975).
[13] Bondi, A.: J. Phys. Chem., 68:441 (1964).
[14] Do Voe, H. et al.: J. Mol. Biol., 4:500 (1962).
[15] 江寿平、乐树云：分子科学学报，2:87 (1982).
[16] Rein, R. et al.: J. Chem. Phys., 42:2177 (1965).
[17] Hankins, D. et al.: J. Chem. Phys., 53:4544 (1970).
[18] Rein, R. et al.: J. Chem. Phys., 41:3393 (1964).
[19] Yamabe, S. I. et al.: J. Amer. Chem. Soc., 97:4458 (1975).
[20] A. 圣-乔其著，林克椿译：«亚分子生物学导论»，科学出版社，(1965).
[21] Mulliken, R. S. et al.: "Molecular Complexes: A Lecture and Reprint Volume", (John Wiley, New York), (1969).
[22] Guryanova, E. N. et al.: "Donor-Acceptor Bond", (John Wiley, New York), (1975).
[23] Flurry, R. L.: J. Phys. Chem., 69:1927 (1965).
[24] 王身立：«控制论与遗传进化问题»，湖南科学技术出版社，(1985).
[25] A. 圣-乔其著，迟翰林译：«电子生物学与癌»，科学出版社，(1979).
[26] R. 霍夫曼，R. B. 伍德沃德著，王志中等译：«轨道对称性守恒»，科学出版社，(1978).
[27] Koshland, D. E.: Proc. Natl. Acad. Sci. U. S. A., 44:98(1958).
[28] Ferreira, R.: J. Theor. Biol., 39:665 (1973).
[29] Racker, E. et al.: J. Biol. Chem., 198:731 (1952).
[30] Kosower, E. M. et al.: J. Amer. Chem. Soc., 78:3493 (1956).
[31] Umeyama, H. et al.: J. Theor. Biol., 46:1 (1974).
[32] Полтев, В. И. И ДР.: Молекулярная биология, 21:704 (1987).
[33] F. W. 普赖斯著，孙冠文等译：«基础分子生物学»，上海科学技术出版社，(1985).

刘次全
谭载友
陈 洪

第4章　核酸的结构与功能

核酸是重要的生物高分子之一，在生命活动中起着极其关键的作用。作为遗传信息的载体，它参与遗传信息在细胞内的传递和表达，从而促成代谢过程及其控制。在一定意义上，可以说生命的一切过程都是根据储存在 DNA（脱氧核糖核酸）内的程序进行的，因此要揭示生命过程的奥秘必须从研究核酸的结构入手，进而探索其生物学功能。关于后者的较详尽讨论，请读者参见第七章——量子遗传学。

4.1　核酸的化学组成和分子结构

作为化学结构的核酸，是由许多称为核苷酸的基本单位组成的生物多聚物。每一个核苷酸又是由三种基本的亚单位：一个含氮碱基，一个戊糖环和一个磷酸基团组成的。核苷酸单体则通过共价的磷酸二酯键相互连接成为长链多聚物（图 4-1）。

如果核酸的糖基是一个五碳糖 D-核糖（图 4-2 *A*），则该核酸称为核糖核酸（简称 RNA）。当核酸的糖基是 D-2 脱氧核糖时（图 4-2 *B*），该核酸称为脱氧核糖核酸（简称 DNA）。

核酸的碱基共有两类：嘌呤碱基和嘧啶碱基。它们的结构和基本骨架的原子标号见图 4-3。从碱基的基本组成原子来看，具有明显的多样性特点，其中包括氢、碳、氮和氧。这种特点决定了核酸碱基具有多方面的反应性能，事实上，在碱基上的确能够发生诸如烷化、卤化、脱氨、氰乙基化和氧的加成等反应。这不仅是储存和传递信息的必要基础，而且也可能是造成突变使物种淘汰和进化，衰老或致癌的原因所在。另一方面，由于碱基种类的多样化，经过适当组合会得出足够多的信息，这是这种生命物质形成信息高分子的可能的和必要的条件。

DNA 和 RNA 的区别除糖基不同外，碱基组成也有差别。在 RNA 中含有尿嘧啶，而 DNA 则含胸腺嘧啶不含尿嘧啶，详见表 4-1。

表 4-1　RNA和DNA的碱基和糖的组分

	RNA	DNA
糖　基	D-核糖	D-2脱氧核糖
嘌呤碱基	腺嘌呤(A)	腺嘌呤(A)
	鸟嘌呤(G)	鸟嘌呤(G)
嘧啶碱基	胞嘧啶(C)	胞嘧啶(C)
	尿嘧啶(U)	胸腺嘧啶(T)

就碱基的种类而言，除了在表 4-1 中所列的 5 种外，在活细胞的核酸中常常还包含一些“稀有碱基”。例如在小麦的胚、细菌和高等生物中发现有 5-甲基胞嘧啶（m^5C），在某些细菌和噬菌

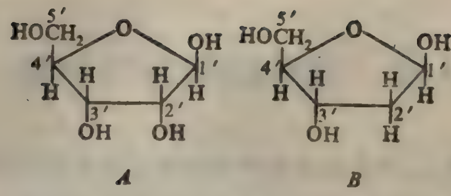


图 4-2 *A*:D-核糖，*B*:D-2 脱氧核糖

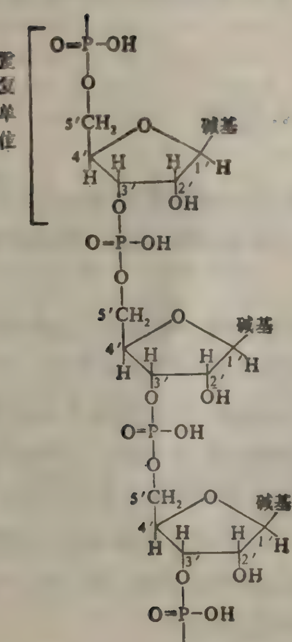


图 4-1 通过单核苷酸之间的磷酸二脂键形成的多聚核苷酸

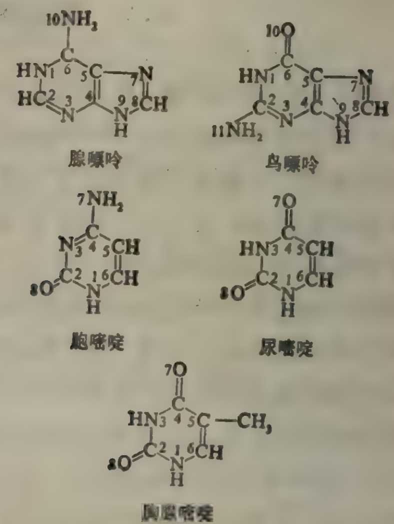
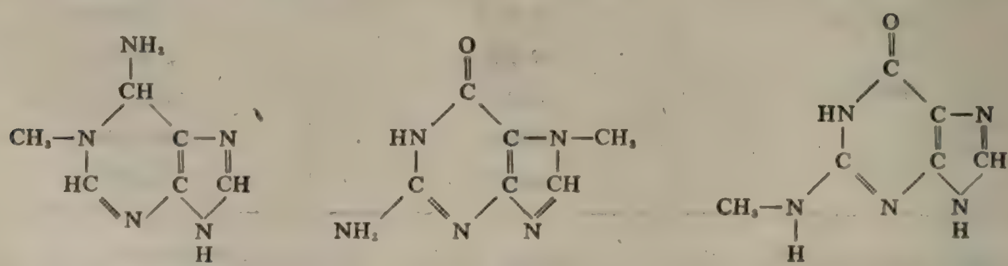
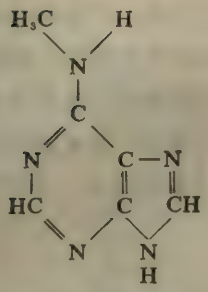


图 4-3 核酸的碱基

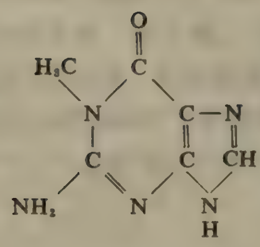
体中则包含有 N^6-甲基腺嘌呤等等。这些“稀有碱基”在生命活动中所起的作用已经引起广泛的重视。图 4-4 列出了部分稀有碱基。



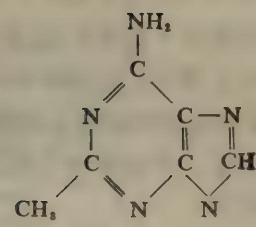
1-甲基腺嘌呤　　7-甲基鸟嘌呤　　N^2-甲基鸟嘌呤



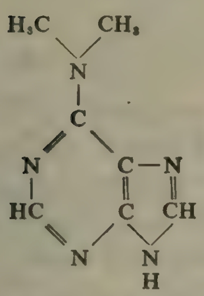
N^6-甲基腺嘌呤



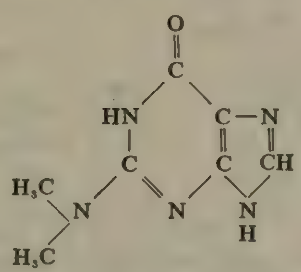
1-甲基鸟嘌呤



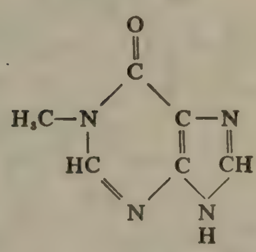
2-甲基腺嘌呤



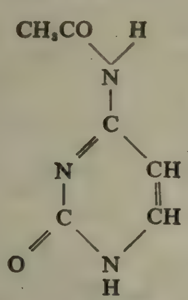
N^6-二甲基腺嘌呤



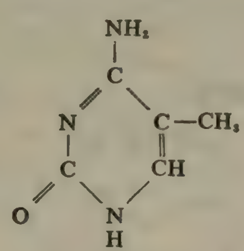
N^2-二甲基鸟嘌呤



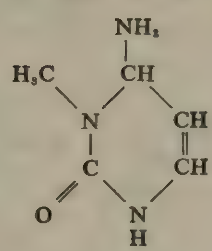
1-甲基次黄嘌呤



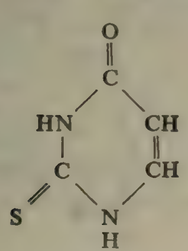
N^4-乙酰胞嘧啶



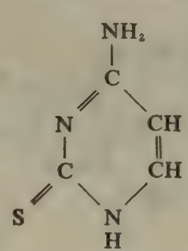
5-甲基胞嘧啶



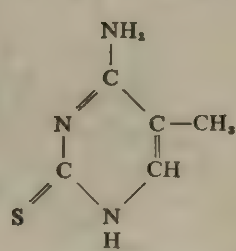
3-甲基胞嘧啶



2-硫尿嘧啶



2-硫胞嘧啶



2-硫-5-甲基胞嘧啶

图 4-4 核酸中的部分稀有碱基

就核苷酸而言，从图 4-5 可以看出，碱基附着在脱氧核糖的第一位碳原子上形成相应的脱氧腺苷、脱氧鸟苷、脱氧胞苷和脱氧胸苷。再连接上磷酸后即形成相应的核苷酸。最后，由核苷酸形成多聚物——脱氧核糖核酸。

脱氧核糖核酸是由两条单链构成的具有双螺旋结构的大分子化合物。在单链结构中脱氧核糖分子与磷酸分子交互排列(参见图 4-1,图 4-6)。两条单链又通过嘌呤与嘧啶之间的氢键使彼此对偶相连。图 4-7 表示 DNA 螺旋的模型图，虚线表示连接碱基的氢键。从图 4-6 中可以看出，两条链由氢键将碱基连接起来：在腺嘌呤与胸腺嘧啶之间有二个氢键，在鸟嘌呤与胞嘧啶之间有三个氢键。从碱基看 DNA 序列，在一条链上是 5′⟶3′，另一条链是 3′⟶5′，因而两条链是反平行的。

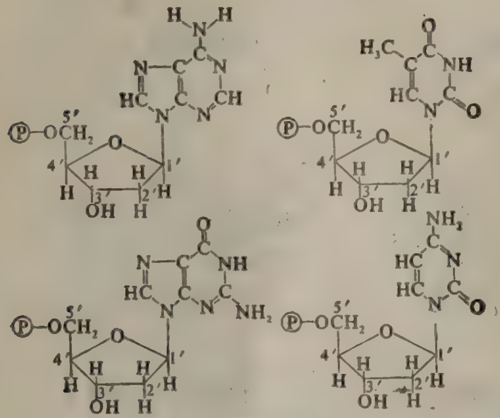


图 4-5 脱氧核苷酸

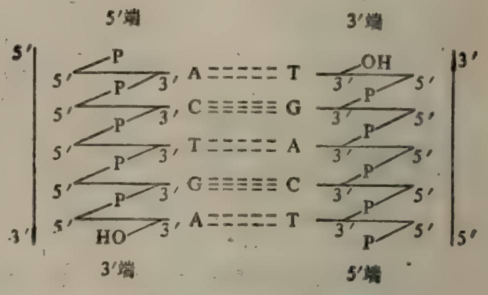


图 4-6

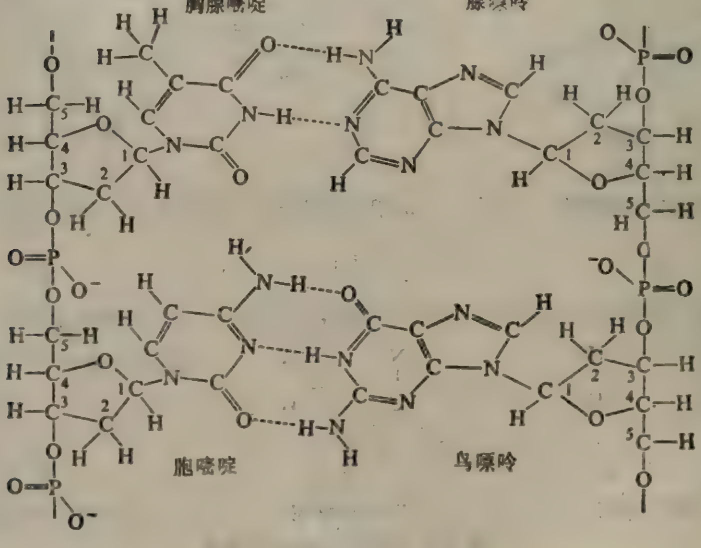


图 4-7

在 DNA 分子的双螺旋结构中，从一个磷酸碱基到下一个磷酸碱基的距离大约是 0.34 nm (指通过上下碱基对平面分子的分子轴的平衡距离)。一个 DNA 分子有点象每圈带有 10 个环的旋转阶梯，每圈的高度约为 3.4 nm，环的半径约为 1.1 nm(图 4-8)。

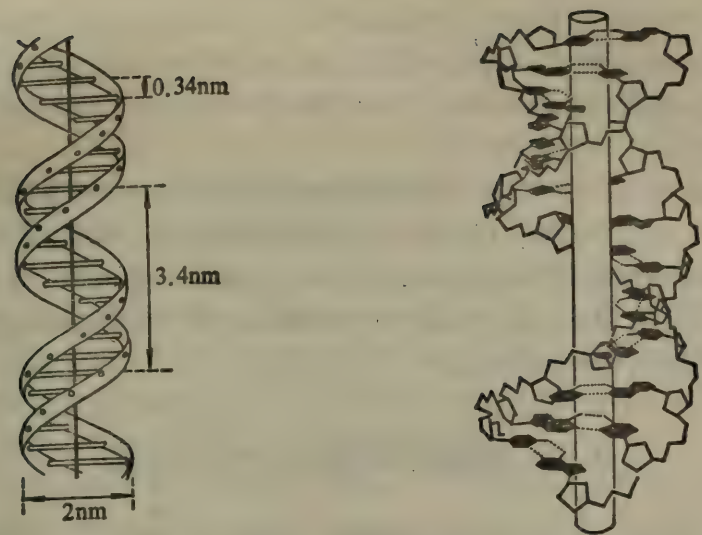


图 4-8 (左)Watson 和 Crick 设计的 DNA 分子示意图。
两条带表示磷酸糖链，横条为连接两条链的碱基对。
(右)DNA 螺旋的模型图，虚线表示连接碱基的氢键。

表 4-2 B-DNA和Z-DNA的结构比较

	B-DNA	Z-DNA
螺旋方向	右手螺旋	左手螺旋
每个螺旋的残基数	10	12(6个二聚体)
直径	～2.0nm	～1.8nm
每个残基上升	0.34nm	0.37nm
螺距	3.4nm	4.5nm
碱基对倾斜	6°	7°
每个残基旋转	36°	−60°(每个二聚体)
糖苷扭转角		
脱氧鸟苷	反式	顺式
脱氧胞苷	反式	反式
糖折迭		
脱氧鸟苷	C2′endo	C3′endo
脱氧胞苷	C2′endo	C2′endo

以上所述是对 B-DNA(右旋的)来讲的，但是几年前又发现了左旋的 Z-DNA，从而使人们对 DNA 结构与功能的认识更加深入，亦更具吸引力。有关 B-DNA 和 Z-DNA 的生物学意义，我们将在以后的有关章节予以简述。在这里，仅列出表 4-2 来初步归纳 B-DNA 和 Z-DNA 在结构上的异同点。

与 DNA 分子不同，RNA 分子一般是单链而不是双链。由于单链分子自身折迭，RNA 分子

中会出现双链部分和单链部分，从而形成某种形式的二级结构。但是，有关 RNA 分子二级结构的研究，尽管运用了物理方法(如小角度 X-射线衍射、圆二色谱、拉曼光谱、核磁共振等)，化学修饰法，部分酶解法和寡核苷酸结合法等等，但迄今为止还没有得到一致公认的二级结构模型。我国科学工作者江寿平等在建立 5 SrRNA、5.8 SrRNA 等 RNA 分子二级结构模型方面，作了一些有意义的工作(参见 11.1.2)[1-3]。

此外，作为规律，对应于 DNA 分子中胸腺嘧啶的位置，在 RNA 中都是尿嘧啶。然而，无论是 DNA 或是 RNA，其线型结构(一级结构) 中存在着的重复单位，是需要我们予以注意的。在 DNA 中，这样的重复单位是指脱氧核糖-磷酸；在 RNA 中是指核糖-磷酸。迄今，有关这样的重复单位在核酸功能中所起的作用的研究，还有若干问题有待进一步深入探讨。在这方面，量子生物学的研究为我们提供了有意义的信息。

4.2 核酸及其组分的电子结构

我们知道，迄今为止，对于 DNA 立体结构的了解要数 B-DNA 了。它的立体结构(在以后的讨论中，凡未加说明的 DNA 分子均系指 B-DNA)已经给出了分子中原子核的空间位置。于是，按照第二章所述，在 Born-Oppenheimer 近似下，原子核和电子可以处理成各自分离的运动，这就为我们提供了进行生物分子电子结构量子力学计算的可能性。但是，正如已经指出过的那样，要用量子力学方法计算一个生物大分子 (哪怕是一个只有 120 个核苷酸的 5 Sr RNA 分子)的电子结构，在实际上是不可能的。即使是对一个完整的核苷酸分子的理论计算，自洽场迭代过程的收敛就是一个问题。然而，值得提出的是，我国科学工作者克服了长期以来不能解决的对稍大一些的分子，甚至小分子在自洽场迭代过程不能收敛的难题，首先成功地应用 ab initio 方法计算了 ATP、cAMP、cGMP 等这样一些分子的电子结构，从而使生物分子的分子轨道研究前进了一步[4-6]。

在量子生物学研究中，由于核酸分子的螺旋对称性和重复单位的存在，因而使我们有可能通过简化和模型化来近似计算一个核苷酸或一小段核苷酸，从而获得核酸分子电子结构的某些信息。

运用量子力学原理和方法来处理 DNA 或RNA 分子的电子结构，首先是进行核酸及其组分分子电子结构的量子力学计算，然后从计算所得到的各种参量来说明这个体系的结构、能量状态及其变化与生物活性和生命过程的关系。当然，如果因此就说能够由纯粹的量子力学，因而由纯粹的物理学原理来预言细胞以至个体的行为和提出全新的生物学定律，那是不太适宜的。尽管如此，它的必要性还是自不必说的。

在实际计算时，一般是依照以下步骤从简到繁一步一步地完成的：

首先，是用量子力学处理单个核苷酸的碱基(A、C、G、T 和 U)的定域 π 电子系，这一步是容易做到的。

第二步，由于 DNA 分子中 A—T、G—C 碱基对是由氢键连接，碱基对的两个部分近似地位于相同的平面内。因此，有可能通过氢键发生 π 电子相互作用。这样，这一步的主要任务

便是在假定存在通常的 π 电子体系相互作用的情况下处理碱基对。实际上，计算的这头两步早已为 Pullman[7]首先完成了。现在，我们可以很容易地用 HMO 法来进行这两步计算。

第三步，这一步要着手考虑沿着同一条链上的堆积碱基之间的不可忽视的 π 电子重叠（π-electron overlap）。为此，可以首先将二聚体作为通常的 π 电子系来处理（这里以及下面将要涉及到的二聚体概念都是指 B-DNA 空间位置中的两个堆积碱基）。而后，假定所有最邻近碱基之间存在 π 电子相互作用，便可采用通常的固体物理学方法获得单个同聚多核苷酸（homo-polynucleotide）链的能级结构，同时还可引伸到计算具有周期结构（如 ACACAC…）的异聚多核苷酸（heteropolynucleotide）的能级结构。再进一步，可以用更复杂的方法处理非周期的一维固体（最邻近的碱基出现率不等），得到其近似的能级结构。

最后一步，同时考虑碱基对之间的氢键和堆积碱基之间的重迭相互作用与Van der Waals相互作用，用微扰法和分子轨道法等近似处理非周期性单调序列双链 DNA 和实际的非周期性的 DNA 分子，得到其近似能级结构，再与实测结果比较。

到目前为止，以上这几步工作已被人们完成了。下面我们就对各步工作所获得的主要结果予以较详尽的叙述。

4.2.1 嘌呤和嘧啶碱基的电子结构[8]

科学研究和教学的实践表明，将嘌呤和嘧啶碱基看作平面共轭体系，有利于对某些问题的理解和实际的计算。但同时，我们也必须清楚地认识到核酸分子结构的复杂性以及其组分间的相互作用和溶液环境的影响等等。因此，在模型的建立和方法的选择等环节上，我们必须注意到这些特点。

4.2.1.1 电荷分布和偶极矩

自 50 年代开始，已经先后应用各种分子轨道法计算了碱基的电荷分布，到 60 年代这些计算达到了高潮。表 4-2 列出了用 EHMO 法、IEHMO（迭代的 EHMO）法、CNDO 法、$\sigma+SCF_\pi$、$\sigma+HMO_\pi$ 以及后来用 ab initio 方法对腺嘌呤、鸟嘌呤、胞嘧啶、尿嘧啶和脑腺嘧啶碱基总净电荷的计算结果。ab initio 法计算结果可参见图 4-9。

从对表 4-2 的分析可以看出，尽管用不同方法所计算出的数值有所不同，但显示出碱基各个原子上净电荷的相对大小基本上是平行的，且碱基的电子结构在许多方面，其主要特征是相同的。如氧、氮原子的净电荷值都是负的（EHMO 法计算的腺嘌呤 N9 为正值，是个例外），而嘌呤碱基的所有碳原子的净电荷值都是正的（用 CNDO 法和 ab initio 法计算的嘌呤 C5 和 ab initio 法计算的鸟嘌呤 C8 是例外），在嘧啶中 C5═C6 键的两个碳原子 C5 和 C6 的原子净电荷值则正好相反（但 ab initio 法计算的 C5 和 C6 原子净电荷值都是负的）。此外，就一般而论，与其它几种方法得到的结果相比较，用 EHMO 法计算得到的杂原子（氮和氧原子）的净电荷值都很大（无论是正的或是负的）。由此可见，该方法跟其它方法之间的差异是比较大的，或许这也正是 EHMO 法固有的缺点。

表 4-3 核酸碱基 U、C、G、A 和 T 的原子净电荷

碱基原子		总净电荷					
		EHMO	IEHMO	CNDO	σ+SCF$_\pi$	σ+HMO$_\pi$	ab initio
U	N 1	−0.291	−0.021	−0.180	−0.175	−0.283	−0.460
	C 2	+1.389	+0.168	+0.447	+0.478	+0.461	+0.484
	N 3	−0.348	−0.038	−0.240	−0.182	−0.324	−0.463
	C 4	+1.157	+0.093	+0.374	+0.370	+0.339	+0.335
	C 5	−0.228	−0.012	−0.171	−0.122	−0.187	−0.078
	C 6	+0.343	+0.056	+0.175	+0.080	+0.192	−0.057
	O 7	−1.408	−0.468	−0.358	−0.451	−0.353	−0.363
	O 8	−1.395	−0.451	−0.373	−0.504	−0.350	−0.406
C	N 1	−0.336	−0.071	−0.175	−0.133	−0.245	−0.463
	C 2	+1.402	+0.055	+0.422	+0.509	+0.455	+0.432
	N 3	−1.024	−0.296	−0.344	−0.584	−0.663	−0.335
	C 4	+0.776	+0.062	+0.324	+0.326	+0.401	+0.240
	C 5	−0.335	−0.049	−0.163	−0.164	−0.155	−0.300
	C 6	+0.290	+0.022	+0.190	+0.113	+0.259	−0.018
	N 7	−0.661	−0.134	−0.222	−0.324	−0.431	−0.598
	O 8	−1.418	−0.519	−0.419	−0.512	−0.388	−0.407
G	N 1	−0.364	−0.055	−0.210	−0.145	−0.254	−0.472
	C 2	+0.975	+0.138	+0.400	+0.418	+0.547	+0.392
	N 3	−1.019	−0.265	−0.373	−0.667	−0.626	−0.361
	C 4	+0.624	+0.067	+0.221	+0.227	+0.308	+0.193
	C 5	+0.117	0	−0.095	+0.098	+0.029	−0.050
	C 6	+1.089	+0.067	+0.355	+0.408	+0.342	+0.359
	N 7	−1.045	−0.283	−0.172	−0.568	−0.540	−0.239
	C 8	+0.613	+0.016	+0.119	+0.214	+0.239	−0.000
	N 9	−0.277	−0.050	−0.103	−0.048	−0.183	−0.447
	O10	−1.409	−0.494	−0.379	−0.493	−0.371	−0.368
	N11	−0.647	−0.102	−0.245	−0.370	−0.417	−0.613
A	N 1	−0.949	−0.277	−0.264	−0.550	−0.518	−0.309
	C 2	+0.636	−0.017	+0.207	+0.369	+0.313	+0.018
	N 3	−0.930	−0.266	−0.257	−0.556	−0.510	−0.307
	C 4	+0.593	+0.070	+0.197	+0.287	+0.334	+0.194
	C 5	+0.154	+0.014	−0.039	+0.098	+0.115	−0.024
	C 6	+0.711	+0.077	+0.266	+0.352	+0.343	+0.234
	N 7	−1.049	−0.277	−0.223	−0.617	−0.557	−0.268
	C 8	+0.618	+0.025	+0.170	+0.250	+0.317	+0.025
	N 9	+0.283	−0.044	−0.123	−0.066	−0.192	−0.456
	N10	−0.648	−0.125	−0.286	−0.361	−0.439	−0.602

Pullman 等对表 4-3 所列六种方法中前四种方法的计算结果作了形象的对 比(见图 4-10、图 4-11)。

表 4-4 总结了五种近似方法计算的四种碱基的偶极矩。

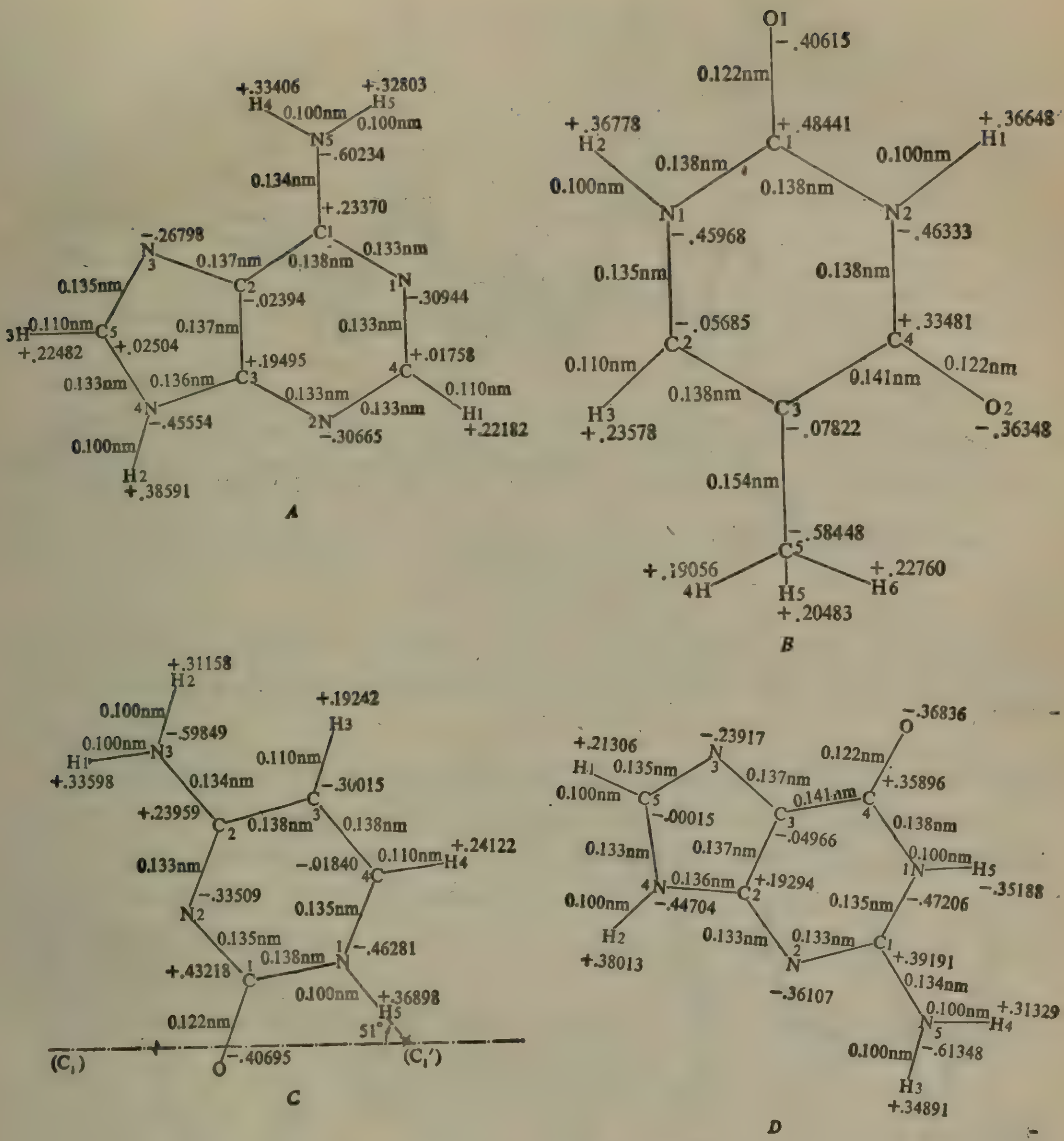


图 4-9 ab initio 方法计算得到的原子净电荷

(*A*) 腺嘌呤, (*B*) 胸腺嘧啶, (*C*) 胞嘧啶, (*D*) 鸟嘌呤[8], *1Å=0.1 nm

从表 4-4 中可以看到，用 EHMO 法计算的偶极矩值最大，这种情况类似于上述用 EHMO 法计算的净电荷值的情况。与 EHMO 法的计算结果相比，用 CNDO 法计算得到的是中等大小的偶极矩值，它跟用包含 σ 分布的 Hückel 法或 PPP 法的计算结果较接近，而且从仅有的实验数据来判断，它们较其它方法都好。从数据上看，鸟嘌呤和胞嘧啶的值大于尿嘧啶的值，最小的值为腺嘌呤。从图 4-12 中可以看出，各种方法所预示的四种碱基的偶极矩的方向也是相似的。

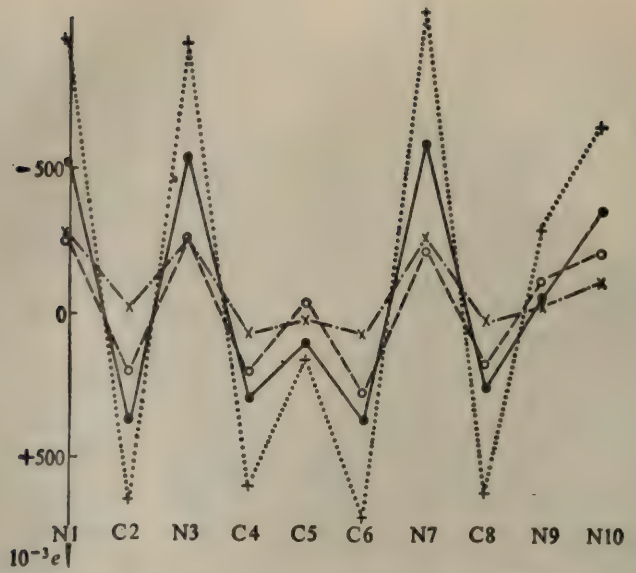


图 4-10 腺嘌呤的总电荷：………，EHT；•—•—·，IEHT；----，CNDO；——，$\sigma+SCF_\pi$.

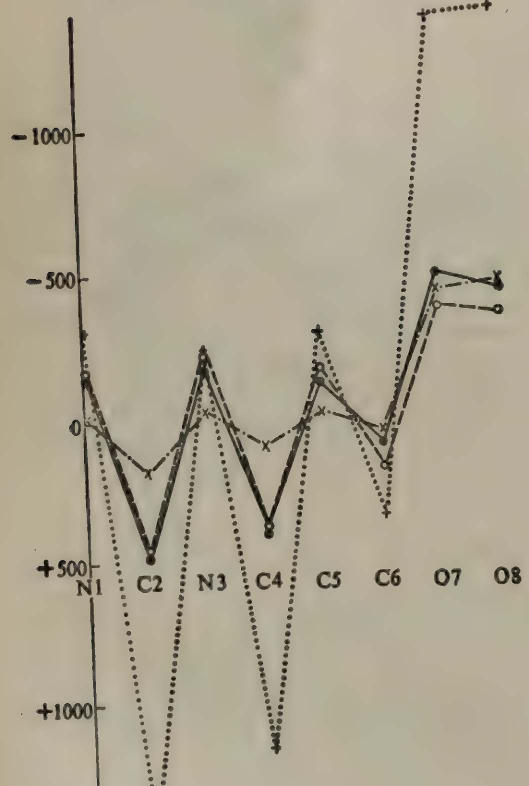


图4-11 胞嘧啶的总电荷：………，EHT；·—·—·，IEHT；----，CNDO；——，$\sigma+SCF_\pi$.

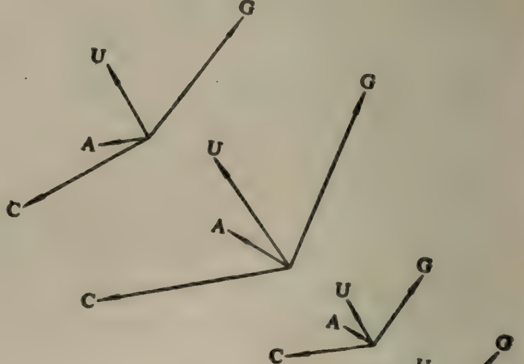


图 4-12 偶极矩的相对大小和方向(自左至右：IEHT，EHMO，CNDO，$\sigma+SCF_\pi$)

表 4-4 四种碱基的偶极矩*

	σ+HMO$_\pi$		σ+PPP$_\pi$		EHMO		IEHMO		CNDO		实验值
腺嘌呤(A)	3.2	77	2.0	73	6.1	65	4.4	104	2.9	64	3.0
尿嘧啶(U)	3.6	39	3.3	35	12.2	41	7.6	36	4.6	36	3.9
鸟嘌呤(G)	6.8	−36	7.2	−35	17.2	−16.5	13.6	−33	7.5	31	—
胞嘧啶(C)	7.2	97	7.1	110	17.3	106	12.9	125	7.6	102	—

* 表中第一列数据是偶极矩，第二列数据是角度(在嘧啶中以N1—C4为轴，在嘌呤中以C4—C5为轴，逆时针测量)。

4.2.1.2 碱性[10]

通常以 pK 值表示杂环的碱性，而且似乎应当平行于嘧啶环上最负的N的 π 电荷 ρ_p^π，然而并未发现这种规律。仅管如此，Pullman 等还是建立了一个反映pK 值与 ρ_p^π 关系的式子：

$$\mathrm{p}K = B + \sum_p \rho_p^\pi r_{np} \tag{4.1}$$

式中，B 为常数，r_{np} 为孤对电子 n 和原子 p 上的 π 电子电荷间的排斥积分。这个关系式反映了各N原子的孤对电子与杂环中所有原子的 π 电子电荷 ρ_p^π 间的相互作用，能正确预测碱性较强的N原子。

(4.1) 式还告诉我们估计杂环上的质子化以及涉及环N孤对电子的其它反应的难易程度，不仅与N原子电荷有关，而且与这个原子上的孤对电子和分子的其它 π电子间的Coulomb 积分值有关。在量子生物学的研究实践中，(4.1)式曾被用于研究若干嘌呤衍生物的抗肿瘤活性与碱性的关系。表 4-5 总结了若干计算结果和由实验得到的 pK_a 值。表中 ρ_q^π 是用 HMO 法得到的相应杂环最碱原子的 π 电子电荷，$\rho_q^{总}$ 是同一原子 q 上的总电荷。数据表明，碱性减少的次序并不与最负N原子电荷减少相一致。

表中列举的 2,6-二氨基嘌呤，6-甲基嘌呤，6-巯基嘌呤，8-氮杂鸟嘌呤等都是具有抗肿瘤活

表 4-5 若干嘌呤代谢拮抗物的碱性

碱基	pK_a	最碱 N	$\sum_p \rho_p^\pi r_{np}$	ρ_q^π	$\rho_q^{总}$	抗肿瘤活性
2,6-二氨基嘌呤	5.1	N 1	−2.21	−0.357	−0.422	+
腺嘌呤	4.2	N 1	−1.91	−0.305	−0.404	天然
鸟嘌呤	3.3	N 7	−1.67	−0.422	−0.346	天然
8-氮杂腺嘌呤	2.6	N 1	−1.65	−0.282	−0.410	—
6-甲基嘌呤	2.6	N 1	−1.40	−0.311	−0.354	+
6-巯基嘌呤	2.5	N 7	−1.42	−0.314	—	+
嘌呤	2.4	N 1	−1.52	−0.308	−0.349	+
8-氮杂嘌呤	2.1	N 1	−1.22	−0.219	−0.351	—
8-氮杂鸟嘌呤	1.1	N 7	−1.46	−0.423	−0.406	+

性的化合物，在结构上类似于天然的嘌呤，因而具有竞争能力，这是抗癌的依据。分析表 4-5 的数据可以看到，显示抗肿瘤活性的化合物通常比无活性的化合物碱性要强，而且活性的抗代谢物往往比无活性的抗代谢物在碱性上更接近天然的碱基。从理论上可以预言最碱的活性代谢物的环氮或是在 N7 或是在N9，这对于腺嘌呤与鸟嘌呤均如此。

4.2.1.3 分子轨道能、电离势和电子亲合势

在表 4-5 中，我们给出了核酸组分中五种碱基的所有占据分子轨道能量，这些数据采用CNDO 法考虑全价电子的计算结果。表 4-7 的数据是由一种 π 电子 SCF 程序获得的五种碱基 π 电子轨道能。表 4-8 则总结了用 HMO 法计算的前线分子轨道（即 HOMO 和 LUMO）能量。实际上，在量子生物学中以 HOMO 和 LUMO 能量作为指数（参见第二章），在预示分子或离子相对的电子给予和接受性质等等方面，取得了跟活性实验相符合的效果。福井谦一在«化学反应与电子轨道»一书中[11]，对“HOMO-LUMO”相互作用等概念作了精辟阐述，读者可从中获得收益。

表 4-6 CNDO法计算的碱基占据轨道能量 *(eV)

U	T	C	G	A
52.87	53.03	52.36	55.49	55.33
45.09	45.87	44.91	51.27	50.39
43.64	43.96	41.69	48.23	44.13
40.22	41.59	38.85	42.73	38.86
36.88	39.61	36.61	39.50	37.34
32.25	32.88	31.66	38.98	33.94
27.41	31.98	27.75	32.08	30.41
π 26.38	27.38	25.74	30.57	29.94
26.36	π 26.90	π 25.53	28.50	27.73
25.71	26.75	25.45	27.82	π 27.07
21.39	24.74	20.96	π 27.60	25.64
π 20.07	π 22.21	20.01	26.48	22.62
19.83	21.35	π 19.91	π 22.82	π 21.15
19.12	20.00	18.40	22.46	20.51
π 18.65	π 19.66	17.91	21.61	19.82
17.72	19.22	π 17.20	20.60	18.07
16.42	18.58	15.56	π 19.66	π 17.89
13.83	π 17.56	13.68	19.55	17.31
π 13.18	17.42	π 13.21	17.38	15.76
12.90	15.88	11.81	16.75	15.11
π 11.88	13.74	π 10.78	15.65	π 14.14
	π 13.10		π 15.34	13.11
	12.85		15.08	π 12.57
	π 11.37		π 14.21	11.46
			π 13.11	π 10.08
			12.77	
			11.55	
			π 9.06	

* 能量值为表中数据的负值。除 π 轨道标明外，其它为 σ 轨道。

表 4-7 PPP-SCF法计算的占据分子轨道能量(eV)(符号相反)

U	T	C	G	A
15.18	15.12	14.23	14.84	14.58
13.12	13.07	12.23	13.68	12.78
12.34	12.39	11.25	11.96	11.32
9.90	10.71	9.33	10.66	9.93
9.15	9.84	8.16	10.09	9.43
	8.82		9.46	7.92
			7.59	

表 4-8 HMO法计算的前线分子轨道能量(β单位)

化　合　物	HOMO	LUMO
腺嘌呤	0.49	−0.87
鸟嘌呤	0.31	−1.05
1-甲基鸟嘌呤	0.30	−1.06
9-甲基鸟嘌呤	0.30	−0.07
胸腺嘧啶	0.51	−0.96
胞嘧啶	0.60	−0.80
5-甲基胞嘧啶	0.53	−0.80
A-T碱基对	0.43	−0.87
G-C碱基对(二个氢键)	0.31	−0.78
G-C碱基对(三个氢键)	0.29	−0.78

从表 4-8 中可以看出，胞嘧啶和 5-甲基胞嘧啶的 HOMO 能量分别为 0.60 和 0.53 β，但 LUMO 能量却完全相同。这对于讨论 5-甲基胞嘧啶的生物学作用是有一定意义的(见 4.9.4)。此表还告诉我们，A-T 碱基对 HOMO 能量低于G-C 碱基对，LUMO 能量高于 G-C 碱基对。

与所有全价电子程序一样，CNDO 法的最显著的特点是表明了 σ 与 π 能级的混合而不出现在 π 电子计算中暗暗假定的表面的"π 壳"。尽管在细节上可能有所不同，但在 π 轨道中间散布 σ 能级这一特点对全电子程序是一致的。

考虑到各种分子轨道法对分子轨道能量计算的差异，我们在表 4-10 中列出了腺嘌呤、胞嘧啶和胸腺嘧啶碱基用 ab initio 法计算得到的分子轨道能量[12]，供读者分析比较时参考。

如果同时再考虑到核酸碱基的电离势和电子亲合势，则更有利于对碱基反应性能的了解。表 4-9 总结了用 EHMO、IEHMO、CNDO、PPP 和 SCF_π 等方法计算五种碱基的电离势值。

表 4-9 电离势 (eV)

化　合　物	EHMO	IEHMO	CNDO	SCF_π	质谱分析实验值
鸟嘌呤	11.84	10.31	9.06	7.59	
腺嘌呤	11.95	10.43	10.08	7.92	8.91
胞嘧啶	12.50	10.39	10.78	8.16	8.90
尿嘧啶	12.71	11.08	11.88	9.15	9.82

表 4-10　ab initio法计算的碱基分子轨道能量(eV)*

C	T	A
565.42	567.00	432.55
429.28	566.78	429.70
428.79	431.87	429.66
428.35	431.07	429.21
315.09	316.69	429.14
314.55	315.07	314.43
313.73	313.28	313.91
311.55	310.87	313.37
39.46	309.81	313.00
36.84	41.37	311.91
35.48	40.18	40.60
32.63	37.80	38.33
31.24	36.00	36.10
26.55	31.77	34.79
25.71	28.11	33.76
23.37	26.91	30.44
22.58	25.95	26.57
21.78	23.02	25.54
20.31	22.58	25.13
19.09	20.99	23.22
π 18.73	π 19.57	21.79
18.51	19.13	20.64
π 16.27	17.72	20.36
16.17	π 17.55	π 19.34
π 14.44	17.52	18.85
11.89	16.80	18.30
π 11.45	π 16.52	17.83
11.40	16.28	π 16.99
π 9.79	15.09	π 14.93
	13.30	14.11
	12.11	12.98
	π 11.91	π 12.79
	π 10.16	11.67
		11.27
		π 9.41
π* 2.20	π* 2.29	π* 2.69
π* 4.60	π* 4.71	π* 3.73
π* 9.77	8.52	π* 5.61
10.26	π* 9.09	9.11
10.80	10.67	9.93
12.35	10.75	10.75
	12.87	π* 10.97

* 未标明的轨道是σ轨道，上面部分包括占据轨道（符号改变），下面部分包括第一虚轨道。

在 CNDO 法中，最高占据分子轨道（HOMO）为 π 轨道，EHMO 法也系如此，但对 IEHMO 法则不然，最高占据轨道为 σ 轨道。这种情况的出现可能是 IEHMO 参数化的人为结果。假如我们认为 π 分子轨道是分子第一电离势的一种测量，那么在 EHMO 和 IEHMO 程序中，能量增加的顺序为 G＜A＜C＜U，而在 CNDO 法中则有 G＜C＜A＜U（见表 4-9，按绝对值比较）。因此，在所有情况中，鸟嘌呤表现出最好的 π 电子给予体性质，而尿嘧啶则是最差的 π 电子给予体。理论计算值与所测电离势相一致：A≈C＜T＜U，其中鸟嘌呤的实验测量虽然尚未完成，但是它具有很强的电子给予体性质已为实验所证实。与实验数据比较，就绝对值而言，全价电子程序计算的结果偏高，而 SCF 法则稍偏低（一般说来 SCF 法计算结果都偏高）。

研究表明，电荷迁移作用的势能与给予体的电离势和接受体的电子亲合势间的差异成正比，亦即，从给予体移去电子或把电子加给接受体越容易，这种相互作用的可能性就越大。事实说明，计算核酸碱基的电离势和电子亲合势，对于了解碱基的电子给予体和接受体性质，进而了解碱基间的相互作用以及碱基与其它分子间相互作用的某些问题将是有益的。

4.2.1.4 电荷密度、自由价、键级和超离域度

在2.2.1.1 中，我们曾经指出：诸如电荷密度、净电荷、自由价、键级和超离域度等等这些理论指数，最初都是在 HMO 法基础上提出的，后又被推广到更精确的计算方案中。

Pullman 等曾经用 HMO 法对核酸碱基的电荷密度、键级和自由价等作过计算[7]。在 4.2.1.1 中，我们比较了包括从头计算法在内的 6 种方法对碱基原子净电荷的计算。迄今为止，对核酸碱基电子结构理论指数的计算已经发表过许多很好的工作，读者可从有关文献中查阅，在这里就不再一一列出。不过，由计算结果归纳出的表 4-11，对我们了解核酸碱基电子结构的理论指数

表 4-11

理论指数	现象
共振能 A的每个 π 电子共振能量大，其次是G、C、U、T。氢键桥的共振稳定性是G-C碱基对大于A-T碱基对。	嘌呤，特别是腺嘌呤(A)对电离辐射和紫外线照射的抗性最强。抗辐射性顺序：A＞G＞C＞U＞T。G-C碱基对较A-T碱基对具有更高的核酸变性温度。
环氮原子的碱性 核酸中最碱性的环氮原子是G中的N 7，在A碱基内的碱性中心是N 3。	核酸的质子化作用和甲基化作用基本上在G的N 7，C的C 5（高等真核生物的甲基化作用均发生在C的C 5）。G的N 7和氧之间最易与金属结合成螯合物，A碱基的质子化、甲基化和N氧化在N 3位置。
氨基中氮原子的电性 NH_2基团中的氮原子一般呈阴性，在A中电荷最多（阳性最小，碱性最强），其次是C和G。	氨基基团的甲基化，在A的NH_2基团中更易与甲醛发生反应，A抗代谢物的抗癌活性最大，其次是C。而无论是A还是G，其亲核部位都不在氨基氮原子上。
氧原子的电性 胞嘧啶的氧原子电负性最强。	胞嘧啶中的氧是不变性DNA质子化作用的基本位置。
非置换碳原子的指数 在G中，C 8位具有最大的自由价，在C中C 5位具有最大的纯负电荷。	这两个碳原子C 8（G）、C 5（C）是进行亲电子置换作用的位置。
C—C键级 最大的C—C键级是T的5—6键，其次是U的5—6键，C的5—6键。	在电离辐射影响下的水合过氧化反应和光分解作用中的水合反应，一般在T的C5—C6键上发生。在DNA中其次是发生在C的C5—C6键。

及其在说明碱基的生物学功能方面是相当有益的，故列出供参考。

作者也曾应用 HMO 法计算 DNA 碱基分子的亲电超离域度。计算结果提示我们，在受到亲电试剂（如致癌性多环芳烃等）攻击的情况下，DNA 碱基分子中容易发出反应的位置是：鸟嘌呤的3位及11位，腺嘌呤的10位，胞嘧啶的7位。在腺嘌呤—胸腺嘧啶碱基对中，容易发生反应的位置是腺嘌呤 N 10 位；在鸟嘌呤—胞嘧啶碱 基 对中则是鸟嘌呤的 N 11 位和C8 位。这些计算结果与文献[13] 报导的实验结果相吻合(参见表 4-12)。

表 4-12 用HMO法计算的DNA碱基分子的亲电超离域度(S_r^E)

DNA碱基分子	位 置	S_r^E
V 腺嘌呤	10	1.9440
	7	1.2628
	3	1.2553
	9	1.2354
	1	1.1959
	2	0.7786
VI 胸腺嘧啶	3	1.6847
	1	1.4946
	7	0.9211
	4	0.8700
	8	0.8654
	10	0.6639
	9	0.5420
VII 鸟嘌呤	3	2.1780
	11	1.9295
	8	1.4980
	1	1.4502
	7	1.2929
	9	1.2585
	10	1.0780
VIII 胞嘧啶	7	1.7578
	1	1.7089
	5	1.4345
	3	1.4222
	8	1.0160
	4	0.7003
IX 腺嘌呤-胸腺嘧啶碱基对	10	2.3265
	11	1.7315
	13	1.6597
	3	1.3083
	7	1.2552
	9	1.2252
	1	1.2240
	8	1.0308
	17	0.9207
	18	0.9056
	14	0.8529
	2	0.7703
	19	0.5420
X 鸟嘌呤-胞嘧啶碱基对	11	1.6886
	8	1.5131
	18	1.5004
	16	1.4033
	15	1.1647
	9	0.8395
	12	0.8077
	19	0.6170
	14	0.5821
	3	0.3429
	1	0.2473
	10	0.1208
	7	0.0356

以上我们讨论了核酸分子的碱基在基态时的若干理论指数，如原子净电荷、偶极矩、分子轨道能、电离势、电子亲合势、自由价、键级和亲电超离域度等，并尽可能与实验值作比较，进而估计各种近似方法(分子轨道法)处理这些基态分子的界限。当然，我们可以容易地看出，即使是像 HMO 这样的简单分子轨道法，在处理某些计算时仍不失为有效的方法。正因为如此，作者认为，在量子生物学的理论计算中，视不同的对象和计算条件，可以采用不同的方法(其中包括简单分子轨道法)，而不必强求采用像从头计算法等精确方法，也可以获得预期的结果。诚然，在必要且具备了运用精确方法的条件时，我们将首先从事这样的计算。

4.2.1.5 碱基激发态的电子结构

在 4.2.1.1 至 4.2.1.4 中，我们比较集中地讨论了碱基在基态时的电子结构的一些问题。下面，我们进而讨论碱基激发态的电子结构。对于碱基激发态的研究，早在 50 年代就已开始。核酸碱基的激发态是指嘌呤和嘧啶碱基在吸收光能后，使整个分子能量提高后的一种能量状态。可是，分子的能量状态又与分子的结构(包括电子结构)及分子的运动形式有着密切的关系。研究生物分子的激发态可以反映它们的构象特点和变化，也可了解光的损伤和失活的详情以及光产物形成等等。

实验研究表明，当用近紫外线照射核酸时，能够布居激发态的只有碱基是唯一的组分分子(因为核糖、脱氧核糖和磷酸根在近紫外线不吸收)。但碱基是杂环结构，对称性很低。可见，能够处理象苯这样的对称分子的分子轨道的一些规则对于了解它们帮助不大，再加之这些发色团彼此又靠得很近，由此能够产生电子相互作用而使问题变得复杂。

人们早已知道核酸碱基受紫外线照射可以发生光化学反应。遗传系统对辐射的敏感性也可以从下述事实得到了解。例如，缺乏正常修补机制的一个细菌，如果它的每一个核酸分子(10^7个核苷酸)形成 10 个光化学缺陷时，就具有 50%致命的机率。这个事件的细节仅仅是由于一个碱基吸收了一个光子而起动，最后在同一碱基或不同碱基上形成一个稳定的光化产物。对此产物只能借助于核酸激发态的有关知识才能探明。

在理论研究方面，量子生物学家已经取得了一些有意义的结果。由于分子轨道法的计算可以得到电子跃迁能等若干有利于揭示核酸碱基激发态的电子结构和特征的指标，因而量子生物学对核酸碱基激发态的研究就具有特定的作用[10]。以下我们简略介绍有关电子跃迁能的计算及其得到的一些结果。

电子跃迁能可表示为式(4.2)的形式，即：

$$\Delta E=\varepsilon_{n^*+1}-\varepsilon_{n^*} \tag{4.2}$$

式中 n^* 表示最高占据轨道。事实上，(4.2)式是碱基的第一激发能的表达式，即是最低空轨道与最高占据轨道间的能量差。

此外，通过分子轨道计算，我们还可以得到跃迁矩：

$$R=\sum_{k=1}^{n}C_{fk}C_{ik}r_k \tag{4.3}$$

式中 C_{fk} 和 C_{ik} 分别表示第 f 和第 i 个分子轨道中第 k 个原子轨道系数，r_k 表示第 k 个核的位置矢量。从(4.3)式中可以看到，在知道碱基的几何结构之后，就很容易计算出由 $i\longrightarrow f$ 的跃迁矩 R。对于 $i=n^*$和 $f=n^*+1$ 的情形，计算得到的跃迁矩 R 见图 4-13。

由图 4-13 可见，嘧啶碱基的跃迁矩近似于平行，嘌呤碱基也是如此。

根据跃迁矩，我们又可以计算振子强度(f)，它表示电子跃迁强度

$$\begin{aligned}f_{fi}&=8\,\pi^2mc\nu\,|R_{fi}|^2 3\,h\\&=1.085\times10^{-5}\tilde{\nu}\,|R_{fi}|^2\end{aligned} \tag{4.4}$$

其中，$\tilde{\nu}=\dfrac{\varepsilon_f-\varepsilon_i}{hc}=\dfrac{1}{\lambda}$为波数。$\lambda$ 为波长，h 为 Plank 常数，c 为光速。

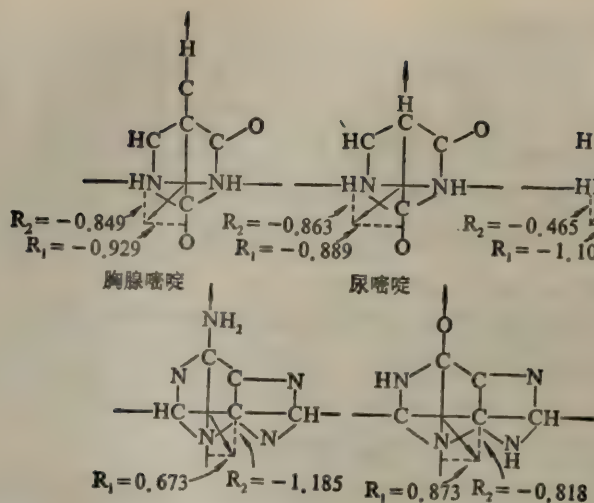


图 4-13 核酸碱基的跃迁矩

振子强度的实验值可根据下式求得

$$f_{实验}=\frac{mc^2}{\pi e^2 N}\int_{\nu_1}^{\nu_2}\delta(\nu)d\nu \qquad (4.5)$$

式中 m 为电子质量，e 为基本电荷，N 为 Avogadro 数，$\delta(\nu)$为克分子消光系数，ν_1 和 ν_2 为频率。

表 4-13 和表 4-14 分别给出了由 EHMO 法和 PPP 法计算所得到的核酸碱基的吸收光谱理论值并和实验值作了比较，其中包括跃迁能和振子强度。

表 4-13 核酸碱基吸收光谱的实验值和计算值(EHMO)

碱基	理论值		实验值	
	nm	f	nm	f
U	259	0.67	259	0.25
T	264	0.67	264	0.21
C	283	0.57	267	0.17
A	276	0.73	262	0.30
G	246	0.63	246	0.40(0.27)

从这两个表中可以看出，无论是那种方法，计算得到的振子强度与实验值都不一致，其中EHMO 法的结果更差。Pullman 等用包括组态相互作用的 SCF 法计算过碱基的激发态，使结果有一定的改进，尤其是第二跃迁位置和带的强度，与实验值符合得较好。但并非所有的 SCF 法都能得到满意的结果，可能的原因在于参数化或其它计算上的缺陷。除了这些工作以外，还有一些人对核酸碱基的激发态作过分子轨道计算。但是，由于在 $n\pi^*$ 和$\pi\pi^*$中有许多低能级，因而

表 4-14 核酸碱基吸收光谱的实验值和计算值(PPP)

碱基	理论值			实验值		
	nm	eV	f	nm	eV	f
U	251	4.948	0.267	260	4.776	0.175
T	253	4.895	0.374	265	4.686	0.183
C	312	3.976	0.102	267	4.642	0.116
A	{273	4.533	0.089	269	4.6	0.049
	{269	4.616	0.288	261	4.758	0.249
G	318	3.895	0.277	275	4.507	0.123

一些工作未能得到可靠的结果。

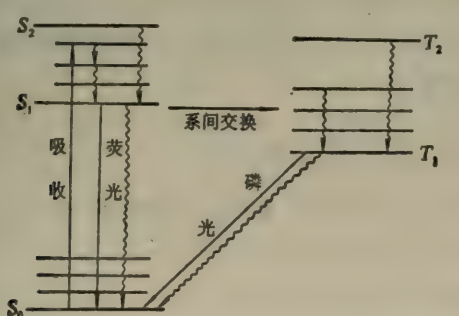
值得指出的是,Savin 等[14]用全价电子近似 CNDO/S,对中性的以及离子型的核酸碱基在基态和较低激发单态和三重态 $\pi\pi^*$ 以及 $n\pi^*$ 的电子跃迁能和电子结构进行了分子轨道计算。计算的结果使得我们能够找出这些分子的最有利的光反应活性位置以及考虑在这些分子的离子态上活性位置定域化的依赖性。同时还指出：中性的嘧啶碱基,其分子片断 C 5—C 6 是最易发生光化学反应的位置,尤其是在 S_1 态的光水合作用和在 T_1 态的光二聚合作用与这一分子片断密切相关*。

如果取在 S_1 态该双键减弱的程度作为光水合有效力的判据,那么尿嘧啶和胸腺嘧啶($\Delta P_{56} = -0.56$)将是最佳的光反应物。在较低的 $n\pi^*$ 态,尿嘧啶和胸腺嘧啶中性分子的电子激发定域在 C 4—O 7 键,Borodavkin 等[15]的工作较详细地讨论了这些问题。

在 S_1 态和 T_1 态的胞嘧啶阴离子,其 C 5—C 6 的键级基本上没有变化($\Delta P_{56} = -0.1$),而 N 1—C 6 的键级表现出明显地减弱($\Delta P_{16} = -0.28$)。当胞嘧啶激发时,片断 N 3—C 4—N 7 的电子结构出现明显的变化。对于嘌呤碱基的离子型而言,一般 C 4—C 5 键在激发时减弱。由于在嘌呤中这个键是整个地被替换,从而可以排除光水合作用类型的反应。另一方面,因为在 T_1 态 C 4、C 5 原子的自旋密度低于嘧啶中 C 5、C 6,故可以推测沿着这个键光二聚合的可能性低。

除以上所述外,还可以举出许多人的工作,如 Maly,Snyder 等[16]用 Gauss 基函数对胸腺嘧啶和阴离子 1-HT⁻与 3-HT⁻的基态和低位激发态进行的全电子的全积分 LCAO SCF 计算。图

*



4-14 则列出了胸腺嘧啶、尿嘧啶和胞嘧啶的基态、激发单线态和激发三重态的 π 键级(用 ASMO SCF CI 法)[17]。

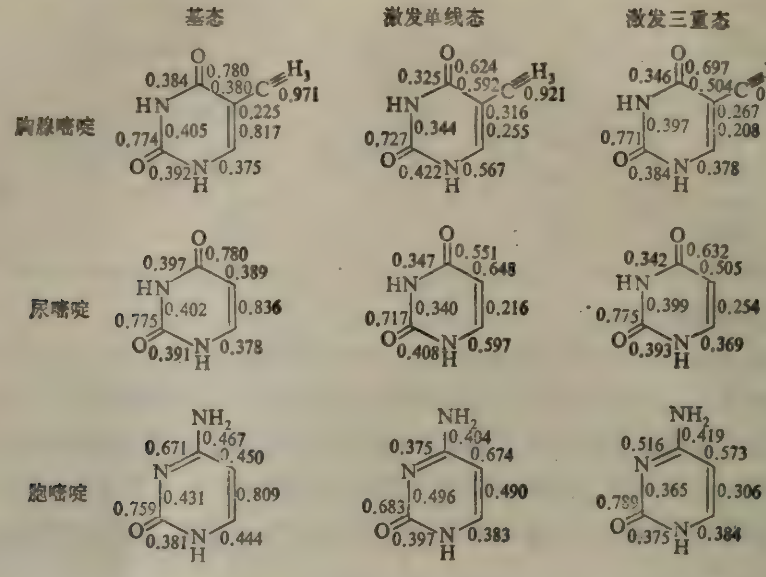


图 4-14 基态、激发单线态、激发三重态的π键级(ASMO, SCF,CI 法)

从图 4-14 中可以看到: 1) 任何嘧啶碱基分子成为激发状态时,5,6 位置的键级较之基态显著为小,容易形成环丁烷型的二聚体。2) 这种倾向明显地表现在胸腺嘧啶是激发三重态,尿嘧啶是激发单线态。3) 胞嘧啶激发状态的 5,6 位置的键级比胸腺嘧啶和尿嘧啶都大,故可解释胞嘧啶二聚体的生成能为何是小的。

如前所述,由于核糖和磷酸根在近紫外线不吸收,因而可以认为碱基的激发态一定具有与相应的核苷酸非常相似的特性。如果碱基与糖结合的位置上是质子化或甲基化(此位置对嘌呤与嘧啶而言,分别是 N 9 和 N 1),且碱基与核苷酸又具有相同的质子化,那么情况确实是这样的。了解这点对我们研究核酸的激发态是有益的。

我们知道,生物分子激发态与能量转移乃是量子生物学中一个发展迅速的研究领域,它不仅是一个理论问题,而且愈来愈成为研究生物分子结构、功能及其相互关系的重要实验方法。

虽然我们对于核酸碱基激发态的了解迄今还没有达到十分理想的程度,但是近几年来,对核酸、蛋白质等生物分子激发态的研究已经十分引人注目,这无疑会促进人们对碱基激发态电子结构的了解,进而逐步认识生物大分子激发态的更多方面。

4.2.2 碱基的互变异构体

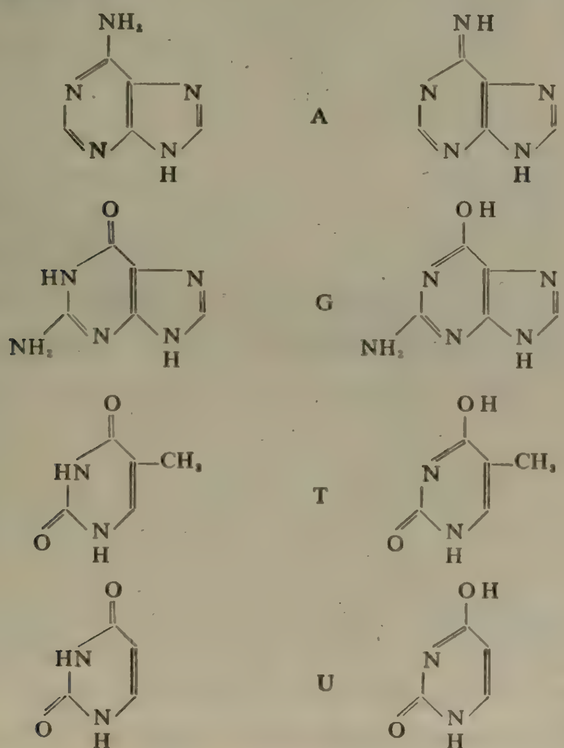
业已知道,在核酸分子中含氧的碱基有鸟嘌呤、胸腺嘧啶、胞嘧啶和尿嘧啶,它们在一定条件下存在着酮式构型和醇式构型的互变。不含氧的腺嘌呤碱基存在着胺型和亚胺型互变。胞嘧啶碱基除了存在酮式和醇式构型互变外,还存在着胺型和亚胺型互变。然而,研究表明,在生理条件下,酮式(内酰胺)是优势结构,含量比醇式(内酰亚胺)要多得多,故一般将前者看作正常型,将

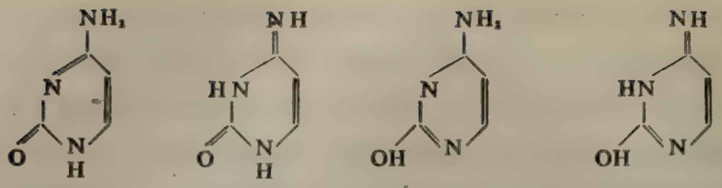
后者看作异常型。但事实说明，当互变异构体和蛋白质、酶或受体结合时，优势结构取何种形式及其量如何，还决定于局部环境。由此看来，所谓正常型和异常型的划分是有条件的，是相对的。

分子遗传学告诉我们，对遗传物质而言，重要的是能够精确无误地复制自身，碱基配对则提供了这种自我复制的手段。因为倘若所形成的两链分子完全有规则，那么碱基配对就保证了新链的顺序对亲本链的顺序是精确无误地互补的。同时，碱基配对也使一条多核苷酸链同另一条多核苷酸链互补顺序之间进行特异性的连接成为可能。一般认为，碱基的互补是由于鸟嘌呤与胞嘧啶、腺嘌呤与胸腺嘧啶（在 RNA 中是鸟嘌呤与胞嘧啶或尿嘧啶，腺嘌呤与尿嘧啶）间存在着氢键的专一性。有关嘌呤与嘧啶碱基晶体 X 射线衍射分析已经证明，当腺嘌呤与胸腺嘧啶通过两个氢键配对，鸟嘌呤与胞嘧啶通过三个氢键配对时，上述碱基的准确配合才会出现。碱基对的这种精确性反映了 DNA 复制的精确无误。

然而，碱基异构体的存在可能对 G-C 和 A-T 配对的精确互补产生不可忽视的影响，从而对上述分子遗传学的一些基本概念和由此而导出的结果提出疑义。在正常细胞中，碱基异构体的存在使已经建立起来的一些有关遗传物质和遗传信息传递的概念和假设变得更加复杂了。至少，从量子生物学的一些研究结果中可以看出这点。

永田亲义曾用 CNDO 法研究过碱基异构体。他的理论计算表明，在 RNA 中，鸟嘌呤和次黄嘌呤的正常型和异常型的分子总能量很接近，而异常型（醇式构型）的分子总能量更低些，碱基 T、C、U 的正常型分子总能量低于异构体分子的总能量（参见图 4-15，表 4-15 和表 4-16）。作者等的计算也得到了类似的结果。





C(I)　　C(II)　　C(III)　　C(IV)

图 4-15　核酸碱基的部分正常型和异常型

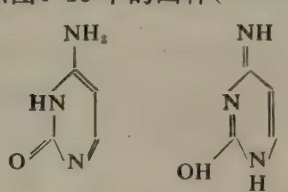
表 4-15　碱基异构体分子总能量的CNDO/2计算

碱　　基	总能量(单位:Hatree)
鸟嘌呤(酮式)	−93.5512
鸟嘌呤(醇式)	−94.0004
胸腺嘧啶(酮式)	−86.3246
胸腺嘧啶(醇式)	−76.3793

表 4-16　碱基异构体共振能的HMO计算(β单位)

碱　　基	共振能	
	酮式	醇式
鸟嘌呤	3.84	4.16
尿嘧啶	1.92	2.14
胸腺嘧啶	2.05	2.27
	胺型	亚胺型
腺嘌呤	3.89	3.62
胞嘧啶　I	2.28	
胞嘧啶　II	2.15	
胞嘧啶　III	2.69	
胞嘧啶　IV	2.41	

由图 4-16 看到，胞嘧啶有四种互变异构体。然而，由于胞嘧啶分子中含有可移动的氢原子，可以形成六种互变异构体，除图4-15 中的四种(I—IV)外，其余两种互变异构体为：

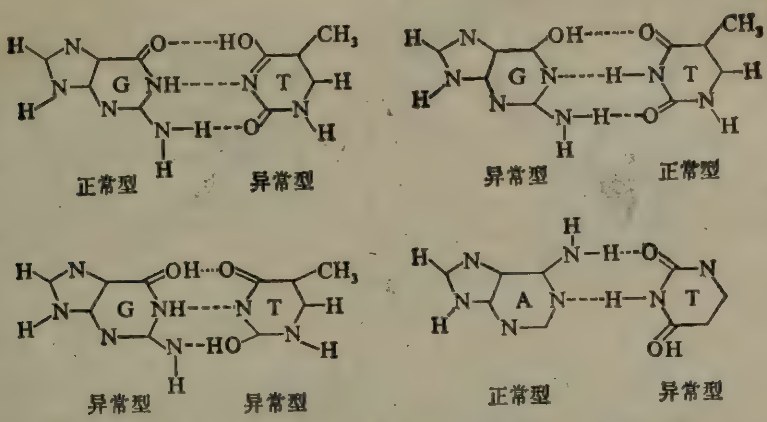


(V)　　(VI)

实验表明，在晶体中胞嘧啶只以(I)式出现。在水溶液中(I)式为主要形式。在理论上，Dewar 用半经验的 MINDO 方法计算了(I、II、III 和V)的生成热，表明(I)式最稳定。而 Goddard 则应用 STO-3G最小基组从头计算研究了(I)至(VI)式的稳定性，表明(III)式最稳定。

Mezey 用ab initio SCF 方法也作了类似的计算[18]。结果表明，对于鸟嘌呤其醇式异构体分

子的总能量低。由于鸟嘌呤和次黄嘌呤的醇式异构体较它们的正常型更稳定，因此可以预料，在鸟嘌呤和次黄嘌呤中会混有相当量的醇式异构体，而这些异构体具有和正常型不同的氢键专一性。由于异构体的存在，我们还可以列出更多的碱基配对方式，例如



等等。但是，在正常细胞中，这些“异常”配对存在的可能性如何？我们仍旧不得而知。不过，根据量子化学计算所提示的结果来看（见表 4-15 和表 4-16），可以推测在正常生理条件下，在 DNA 的双链结构(或RNA 单链因回折而形成的双链结构)中，可能存在着鸟嘌呤的醇式构型与胸腺嘧啶的酮式构型通过三个氢键的专一配对。如果事实果真如此，那么这样的碱基配对的存在将会影响到碱基配对的性质，从而可能导致 DNA 复制和转录的错误。

4.2.3 碱基的分子静电势[19]

通过对核酸碱基的分子静电势计算，我们可以获得那些信息呢？从图 4-16 中可以看出，这样的计算可以很好地解释核酸碱基对于许多亲电试剂(如质子、烷化剂等等)的反应活性和这些碱基间势能极小值的差别。

从图 4-16 中，可以明显地看到，鸟嘌呤环上的 N 7 位置（－91 kcal/mol），胞嘧啶的 O8 位置（－83.1 kcal/mol）、腺嘌呤环上的 N 7 、N 3 位置（－66kcal/mol 和－71 kcal/mol)、尿嘧啶的 O 4 位置(－57.8 kcal/mol)，在反应中都是较活泼的位置。

4.2.4 核酸碱基的配对[19]

碱基配对问题是核酸电子结构中研究得较多的一个问题。在 4.2.2 中我们已经提及核酸碱基相互间的配对,对于 DNA 的自身复制以及转录和 RNA 转译的重要性。

在这一节里，我们将着重讨论碱基配对时，两种碱基分子是如何配对的？能否找到形成这些碱基对的力的作用以及对这些力的作用的解释？这正是我们力图回答的问题。

在 4.2.1.1 里，讨论并给出了计算的腺嘌呤、鸟嘌呤、胞嘧啶和尿嘧啶偶极矩的相对大小和方向。由于分子中正电荷分布中心与负电荷分布中心并不总是重合的，这样就产生了一

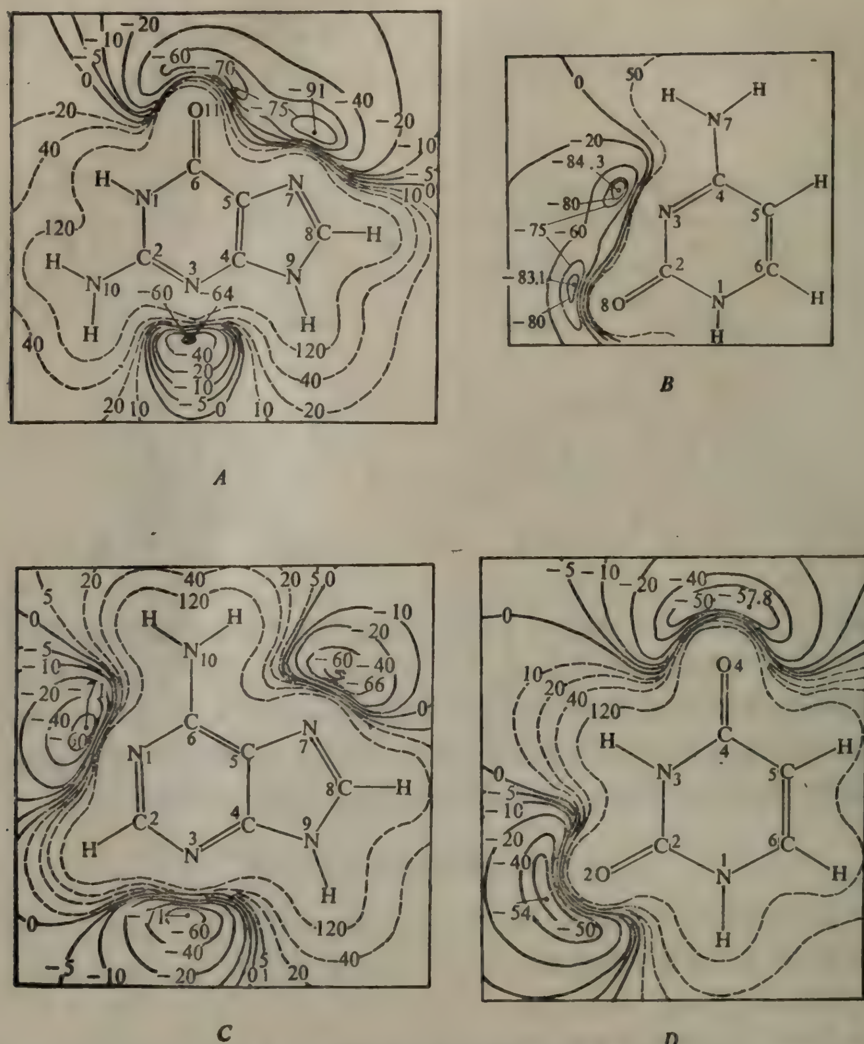


图 4-16　核酸碱基在平面上的等静电势线图：
(*A*) 鸟嘌呤，(*B*) 胞嘧啶，(*C*) 腺嘌呤，(*D*) 尿嘧啶
(单位：kcal/mol，实线表示负势，虚线表示正势)。

对偶极子及偶极子间的相互作用。实验和计算结果指出，碱基对的偶极矩是显著的，碱基对之间的力比起沿螺旋而堆积的碱基之间的力具有更大的极性相互作用。

图 4-17 给出了由 Pullman计算得到的偶极矩与碱基的关系，有助于说明腺嘌呤与尿嘧啶、鸟嘌呤与胞嘧啶相互配对的事实。

一般说来，计算碱基对之间的偶极相互作用有两种方法。一种方法是把相互作用能当作三种不同来源的总贡献，即偶极—偶极相互作用，偶极—诱导偶极相互作用和 London 力。

若以 μ 和d 分别表示偶极矩与极化率，当距离是 R，电离势为 I，则 A 与 B 两个碱基之间的偶极相互作用能的各项贡献可写作：

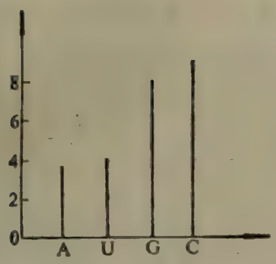


图 4-17 碱基与偶极矩的关系

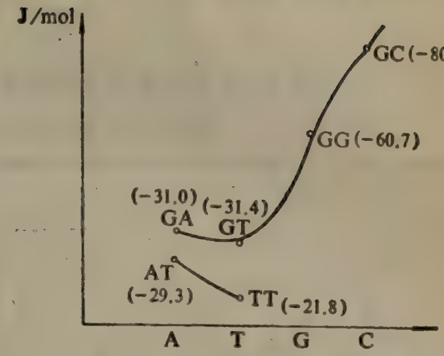


图 4-18 DNA 碱基间的结合能

$$E\mu\mu=\frac{1}{R^3}\left[\mu_A\cdot\mu_B-\frac{3}{R^2}(\mu_A\cdot R)(\mu_B\cdot R)\right] \tag{4.6}$$

$$E\mu\alpha=-\frac{1}{2}\frac{1}{R^6}\left\{\alpha_A\left[3\left(\frac{\mu_B\cdot R}{R}\right)+1\right]+\alpha_B\left[3\left(\frac{\mu_A\cdot R}{R}\right)+1\right]\right\} \tag{4.7}$$

$$E_L=-\frac{3}{2}\frac{I_A I_B}{I_A+I_B}\frac{\alpha_A\alpha_B}{R^6} \tag{4.8}$$

第二种方法是将碱基对之间的偶极相互作用作为由分开的正电荷与负电荷之间的相互作用再加上电荷诱导偶极相互作用与 London 力的贡献。其中 London 能与(4.8)式相同，其它两项贡献为：

$$E\rho\rho=\sum_{i\in A}\sum_{j\in B}\frac{\rho_i\rho_j}{R_{ij}} \tag{4.9}$$

$$E\rho\alpha=-\frac{1}{2}\left[\mu_{AB}\cdot E_{AB}+\mu_{BA}\cdot E_{BA}\right]$$

$$=-\frac{1}{2}\left[\alpha_B(E_{AB})^2+\alpha_A(E_{BA})^2\right] \tag{4.10}$$

其中，ρ_i、ρ_j分别是碱基 A、B 中原子的净电荷。μ_{AB}是 A 碱基上的电荷作用在 B 碱基上的诱导偶极；E_{AB}是 A 上的电荷作用在 B 上引起的电场。$\mu_{BA}E_{BA}$ 有同样意义。

通过 3.2 节对分子间相互作用的“偶极”近似和“单极”近似的讨论，我们不难理解(4.6)至(4.10)式的意义。在两种常用的计算碱基对之间的偶极相互作用的方法中，已有研究表明，第二种方法较第一种方法给出的结果更精确。Pullman 用第二种方法计算出了碱基对的各种结合的能量，指出能量的最低处是发生氢键的结合处。能量越负，碱基间的结合越紧密，越有利于相互作用。计算结果见图 4-18。

从图中可以看到，在两条曲线中结合能最负的也即是最强的结合。每条曲线上最稳定的碱基对都是迄今为止已知实际存在于 DNA 双螺旋中的，即 G-C(−19.2 kcal/mol) 和 A-T (−7.0 kcal/mol)。实际观察 DNA 碱基对的结果与计算结果符合得很好。不仅如此，计算结果还显示出 G-C 对的稳定性(负能值)要比 A-T 对大，这有助于说明凡是含 G-C 对较多的 DNA，其热变性温度也较高这一事实。然而，我们也不得不注意到，除了 G-C、A-T 等配对外，还有

可能出现所谓“反常配对”以及在 RNA 回折结构和 DNA 在“转录”过程中的G-U、A-U 配对等。

表 4-17 和表 4-18 分别给出了各种可能的A·U配对和G·C配对的相互作用能[8]。

表 4-17　各种可能的A·U对的相互作用能(kcal/mol)

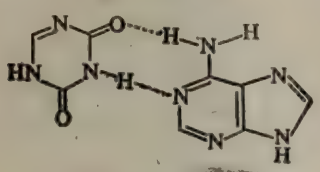
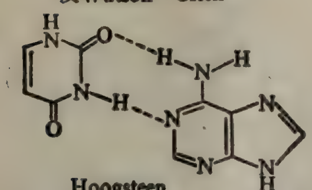
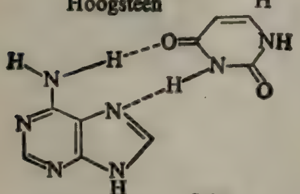
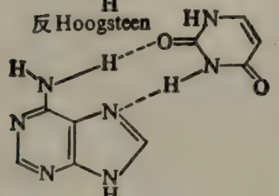
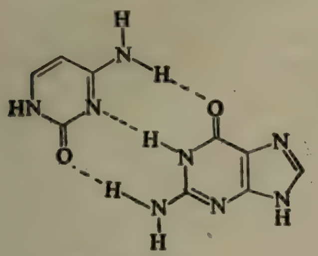
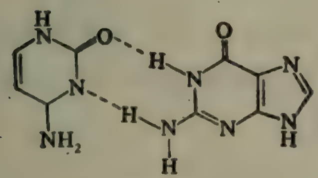
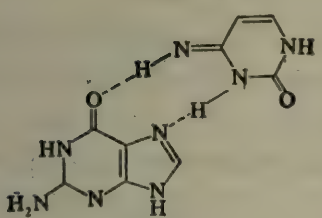
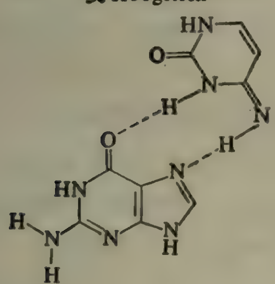
构　　型	E_{pp}	E_{pq}	E_L	$E_总$
Watson–Crick 	−4.64	−0.25	−0.69	−5.58
反Watson–Crick 	−3.98	−0.25	−0.71	−4.94
Hoogsteen 	−5.86	−0.22	−0.88	−6.96
反Hoogsteen 	−5.63	−0.17	−0.94	−6.74

表 4-17 告诉了我们有趣的计算结果，即 Hoogstee 及反 Hoogsteen 构型的A·U对，其 $E_总$ 明显低于 Watson-Crick 及反 Watson-Crick 构型的A·U对的相互作用能，而G·C对的相互作用能，则以 Watson-Crick 构型为最低(见表 4-18)。由此可见，A·U 配对的最稳定构型是 Hoogsteen 型； G·C配对的最稳定构型是 Watson-Crick 型。这些结果与共晶现象中的观察是一致的。

由表 4-19 的数据可以进一步看出，就 Watson-Crick 意义上的碱基互补专一性的理论依据。

表中第一列数据相应于各种碱基氢键自身缔合相互作用能的极大值，第二列相应于交叉缔合的碱基相互作用能。当然，这些计算包括了一些可能是并不存在的缔合形式。就自身缔合而言，G·G和C·C是相对较强的，而A·A与T·T是较弱的。对于交叉缔合，A·T和G·C作

表 4-18　各种可能的G·C对的相互作用能(kcal/mol)

构　　型	E_{pp}	E_{pa}	E_L	$E_{总}$
Watson－Crick 	−15.91	−2.02	−1.25	−19.18
反Watson－Crick 	−10.24	−1.99	−0.62	−12.85
Hoogsteen 	−3.98	−1.33	−0.44	−5.75
反 Hoogsteen 	−2.42	−0.91	−0.34	−3.67

用能分别大于它们两组分自身缔合作用能。

现在回到这一节开始时，我们力图想要回答的问题上进行讨论。其实，迄今为止，我们仍旧没有能够从量子生物学的角度，就核酸碱基的配对问题作出令人满意的确切回答。但是，从两个碱基之间的偶极相互作用入手，通过理论计算已经获得了与实验值和实际观察在若干方面相符合的结果。不仅如此，我们还介绍了有关碱基异构体和非 Watson-Crick 碱基配对的量子化

表 4-19　形成氢键的碱基对的相互作用能 (kcal/mol)

A·A	−5.8		A·T	−7.0
T·T	−5.2			
G·G	−14.5		G·C	−19.2
C·C	−13			
A·A	−5.8		A·C	−7.8
C·C	−13			
G·G	−14.5		G·T	−7.4
T·T	−5.2			
C·C	−13		C·T	−6.5
T·T	−5.2			
A·A	−5.8		A·G	−7.5
G·G	−14.5			
	A·T	>	A·A	
			T·T	
	G·C	>	G·G	
			C·C	
	A·C	<	C·C	
	G·T	<	G·G	
	C·T	<	C·C	
	A·G	<	G·G	

学计算。然而,对于那些在活细胞 DNA 中,或在由 DNA ⟶ RNA ⟶ 蛋白质的遗传信息传递过程中实际存在的某些碱基配对，则还没有或者还无法作出较为确切的回答。有关碱基间的相互识别，读者可通过第三章有关内容的回顾而加以了解。

4.2.5　多聚核苷酸

在这一节里,我们试图遵循这样的途径,即从核酸的基本组分——碱基、核糖或脱氧核糖和磷酸开始,经过核苷 ⟶ 核苷酸 ⟶ 多聚核苷酸,最后讨论 DNA 或 RNA分子的量子化学计算,从而探讨核酸的结构与功能,这一揽子的、十分困难的问题。对于第一部分，在 4.2.1 至 4.2.4 里已经作了讨论，至于其中的核糖或脱氧核糖和磷酸将在第二部分的叙述中予以讨论。

4.2.5.1　核苷

对于核苷的电子结构已经作过一些研究。这里，我们仅以 Kwiatkowski 和 Pullman 等的工作为例作一简述。他们首先运用从头计算法对核苷的电子结构作了计算[20]，其中腺苷净原子电荷分布的计算结果如图 4-19 所示。

从图中可以看出,腺嘌呤碱基与核糖结合成为腺苷后,其净电子电荷除腺嘌呤的 N 7 位和核糖的 C 1′位外,其余位置无明显改变。从N原子电荷分布看,在腺嘌呤和顺式-腺苷中为 N 1>N 3>N 7,在反式-腺苷中则为 N 3>N 1>N 7,而核糖的电荷在图 4-19 所示的几种情形里均为 O 2′>O 5′>O 3′>O 1′,这些结果向我们提示了腺苷在参与反应时的若干反应部位。

4.2.5.2　核苷酸

我们知道,核酸是由核苷酸组成的多聚物。因而核苷酸的量子生物学研究,就成了探索核酸电子结构的基础。

现在,我们以腺苷三磷酸(ATP)为例(分子结构如图 4-20 所示),试看通过分子轨道研究,对其电子结构与功能的认识上，获得了哪些有益的信息。

腺苷三磷酸分子是大家熟知的重要生物分子(系核酸组分核苷酸的结构类似物),它在有机体代谢中起着重要的作用,是许多生理过程的主要能源。当 ATP 水解成腺苷二磷酸(ADP)时,

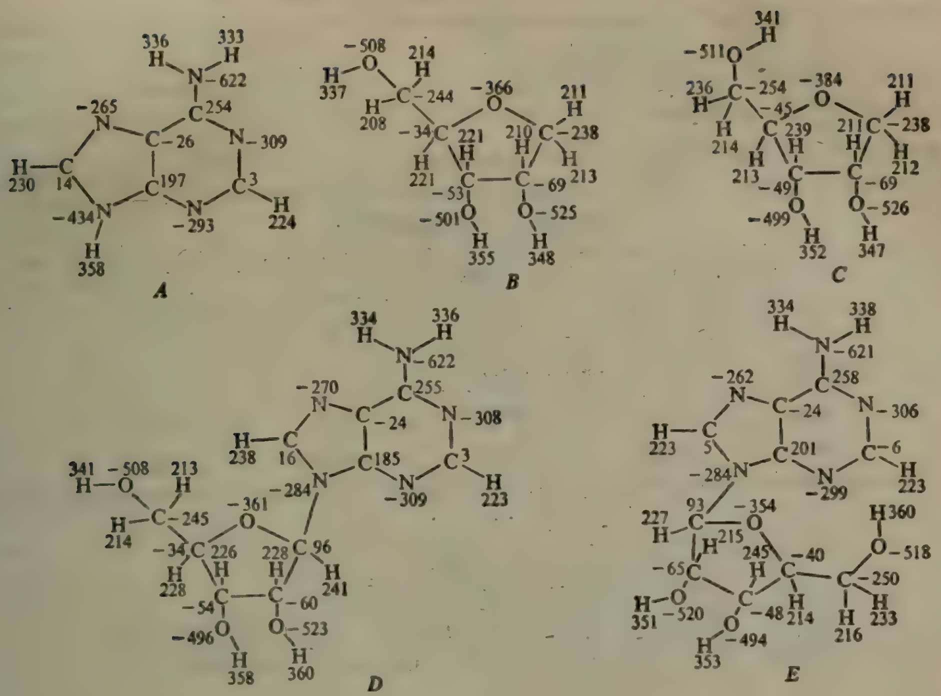


图 4-19 腺苷净原子电荷分布(10^{-3}e)。

(*A*) 腺嘌呤,(*B*) gt 3′-端核糖,(*C*) gg 3′-端核糖,(*D*) 反式-腺苷和(*E*)顺式-腺苷

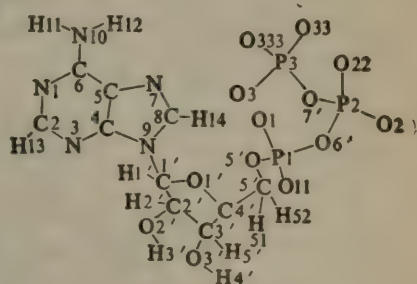


图 4-20 ATP分子结构

自由能大约变化−9 kcal/mol 左右。由于这个水解过程放出这样大的自由能是随着 ATP 的 P—O 键切断而产生的,故称此键为“高能”磷酸键。为了阐明它的生物学功能,早在 1960 年就有人应用HMO法对其进行研究[7,17]。后来,Boyd 和 Lipscomb 用 EHMO 方法对此分子进行了详细的研究[21]。为了得到更好的结果,一些人试图通过用 CNDO 法对 ATP 分子进行研究[17,22],但均因计算比焦磷酸大的多磷酸时不能收敛而失败。我国科学工作者江寿平等解决了不收敛这个困难问题,并用 CNDO法和 ab initio 法计算了 ATP 和 ADP、AMP 等分子[4-6,23]。

表 4-20—4-22 分别给出了 ATP、ADP 和 AMP 分子的总能量、本征值和偶极矩以及原子净电荷。为了便于比较,表 4-21 中也列出了 Boyd 和 Lipscomb[21], Alving 和 Laki[22]等人的计算结果。而 ATP 分子磷酸链部分的电荷分布则示于图 4-21 中,它的键级表示于图 4-22 中。

由表 4-20—4-22 和图 4-21、4-22 可以归纳出:(1) 在 ATP、ADP 和 AMP 分子中,腺嘌呤部分同分子的其它部分的相互作用非常小。(2) 磷酸链上的原子电荷分布是正负交替的。于是,一直认为“高能”键是由于磷酸链内存在静电排斥的结果的说法是不成立的。(3) ATP、ADP

表 4-20　ATP,ADP和AMP分子的总能量(a.u)

分　子	ATP		ADP	AMP
方　法	从头计算法	CNDO	CNDO	CNDO
总 能 量	−2545.92676	−398.81090	−337.06186	−274.87254

表 4-21　本征值和偶极矩

分　子	最高占据轨道(a. u)		最低空轨道(a. u)		偶极矩(D)	
	CNDO	从头计算法	CNDO	从头计算法	CNDO	从头计算法
AMP	−0.3636		0.1091		10.3810	
ADP	−0.3626		0.0732		12.9601	
ATP	−0.3580	−0.1745	0.0790	0.2858	11.2122	9.2990

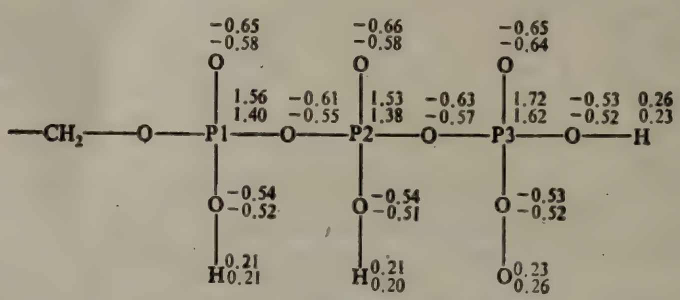


图 4-21　磷酸链上原子的净电荷
(上为 CNDO 结果,下为从头计算法结果)

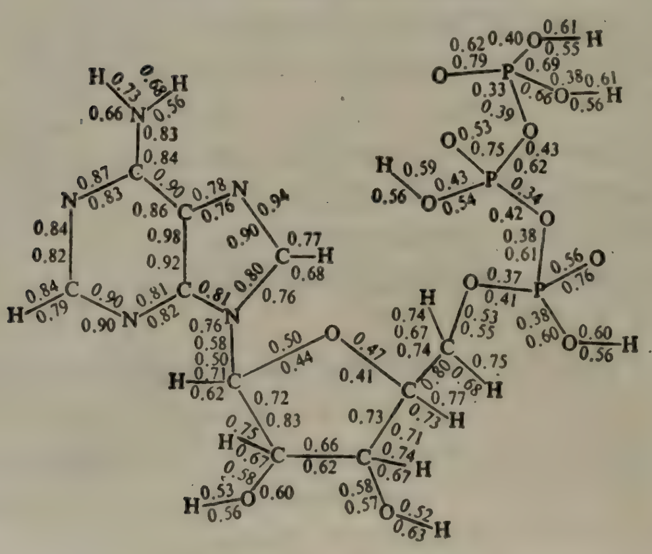


图 4-22　ATP 分子的键级
(上为从头计算法结果，下为 CNDO 结果)

表 4-22 ATP, ADP和AMP的原子净电荷

分　　子	AMP		ADP	ATP		
方　　法	CNDO	CNDO[21]	CNDO	CNDO	EHMO[21]	从头计算法
原子						
N 1	−0.30	−0.29	−0.31	−0.31	−0.86	−0.27
C 2	0.19	0.20	0.19	0.19	0.51	0.09
N 3	−0.23	−0.25	−0.23	−0.22	−0.76	−0.25
C 4	0.22	0.19	0.23	0.23	0.45	0.16
C 5	−0.07	−0.05	−0.08	−0.08	0.08	−0.01
C 6	0.24	0.26	0.24	0.24	0.60	0.23
N 7	−0.24	−0.18	−0.25	−0.27	−0.83	−0.26
C 8	0.16	0.13	0.16	0.17	0.40	0.14
N 9	−0.07	−0.14	−0.06	−0.06	−0.10	−0.19
N10	−0.21	−0.21	−0.21	−0.21	−0.34	−0.33
H11	0.11	0.11	0.11	0.11	0.21	0.18
H12	0.11	0.10	0.11	0.10	0.24	0.17
H13	−0.01	−0.01	−0.01	−0.01	0.08	0.05
H14	−0.01	0.02	−0.01	−0.01	0.09	0.06
C1′	0.19	0.22	0.19	0.20	0.59	0.12
C2′	0.12	0.10	0.12	0.12	0.36	0.05
C3′	0.13	0.13	0.13	0.13	0.38	0.07
C4′	0.10	0.12	0.10	0.10	0.39	−0.02
C5′	0.16	0.14	0.16	0.16	0.32	0.00
O1′	−0.23	−0.24	−0.23	−0.24	−0.85	−0.07
O2′	−0.26	−0.26	−0.26	−0.26	−0.83	−0.24
O3′	−0.29	−0.28	−0.30	−0.29	−0.84	−0.30
O5′	−0.43	−0.25	−0.42	−0.42	−0.39	−0.32
H1′	−0.01	0.00	−0.01	−0.02	0.05	0.05
H2′	0.00	0.00	0.00	0.01	0.04	0.06
H3′	0.15	0.15	0.15	0.15	0.43	0.16
H4′	0.17	0.16	0.17	0.16	0.44	0.19
H5′	0.00	0.00	0.00	0.02	0.03	0.06
H51	−0.02		−0.03	−0.03	0.07	0.03
H52	−0.05		−0.05	−0.05	0.07	0.02

和 AMP 分子的最高占据分子轨道(HOMO)的本征值是相同的，也即它们的电离势也相同，在化学反应中具有相近的给电子能力。(4) 在磷酸链周围的氧原子带有较负的电荷。因而有很大静电排斥力，这就有利于磷酸键的水解断裂。

鉴于 ATP 分子在生命系统中的重要作用，以下我们将围绕着对它所进行的分子轨道研究，讨论几个问题。

1) 高能磷酸键的本质和电子结构

(1) 静电排斥　Pullman 等[24]曾强烈主张：P—O—P 键由于正电荷相邻排列着，所以它们相互排斥发生键的不稳定。但是，在 P—O—P 上排列着正电荷，是因为 Hückel 法基于仅考虑 π 电子的限制(图 4-23)，如果按照考虑全价电子的推广的Hückel 法和CNDO法、从头计算法

(见图 4-21)的计算结果，就不出现相邻的正电荷电子结构，而是正负电荷交替的电子 结构。 因此，ATP 的高能键的本质，基于在 P—O—P 上存在相邻的正电荷相互排斥，这一想法当然被否定。

(2) ATP 是不是高能键? 这个问题在文献[17]中已作过简单讨论，但没 有获得明确的结论。通过 CNDO 法和从头计算法的计算，使我们有可能在更精确的计算结果的基础 上 来讨论这一问题。

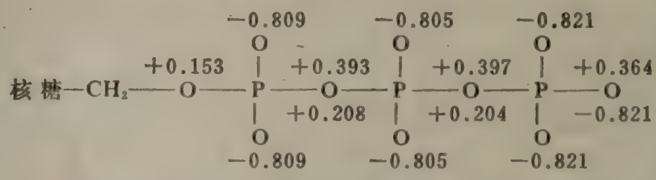


图 4-23 3Hückcl MO 法计算得到的 ATP 的电荷分布

高能磷酸键是指水解自由能，而不是能量贮存在P—O键。如前所述，当 ATP 水解成ADP 时， 自由能大约变化−9 kcal/mol 左右。 早期那种以磷酸键骨架内存在静电排斥的假设来说明 “高能”键的存在是不成立的。 但在磷酸键周围的氧原子带有较负的电荷，因而相互排斥，这是有利于磷酸键水解断裂的。至于 ATP 的水解位置，按江寿平等的计算，键级是键能的表示，它可比较同类型的键[25]，可提供键强度的倾向。在 ATP 分子的磷酸链上，键级最小的即意味着 这 个键最弱，最易断裂。由图4-22可见，磷酸链上键级最小的是O 7′—P 3键(0.39)，此键在水解时 趋 向于最易断裂，实验研究也证实 ATP 分子水解时正是此键断裂[17]。

2) 在生物体反应中 ATP 的专一性

一般 认为，高能键的化合物，除 ATP 外，还有 ADP、磷酸烯醇丙酮酸、乙酰 磷 酸 等，虽然是很多，但为何在生物体内只有 ATP 被作为能源利用呢? 这一问题至今尚未得到充分 的 解答。

与这一问题相联系而被注意的是生物体内 ATP 的利用，为何离不开二 价 金 属 离 子 (如 Mg^{2+})的存在? Szent-Gyŏrgyi 认为[26]，ATP和 Mg^{2+} 离子生成的配合物帮助了 ATP 在生物体内专一性的利用。实验上也得到了支持这种想法的结果。例如，按照 EPP 等[27]的结果，从ATP 与 Mg^{2+} 的配合物的红外吸收检测了 Mg^{2+}和腺嘌呤或磷酸等部分的相互作用。但是，这种相互作用是存在于分子内还是分子间尚未肯定，故明确的结合体结构也未能确定。在此意义上，考虑各种可能的配合物模型，并且研究模型的妥当性，无疑是有意义的。

迄今，已经用核磁共振，拉曼光谱及其它实验方法对金属离子和 ATP 结合 进行了广泛的研究。由表4-21可见，磷酸链上 O 1，O 2和O 3 是很负的原子，因而它们和金属离子有很强的相互作用。所以当 ATP 分子处于折叠构象时，二价的金属离子是可能和 O 1，O 2和O 3又和碱基 部 分 N 7 等原子相结合。这个 ATP 和二价金属离子形成的配合物(ATP—M^{2+})在一些酶反应中，是作为底物结合在酶分子的活性中心的。而且 ATP 分子碱基部分的 N 1 原子可能和酶分 子中 带正电的基团产生静电的或氢键键合的作用。这和 N 1 附近基团有关。一个可能的结合模型示于图4-24中，这些结合位置和实验结果是一致的[28]。

总之，ATP 和 M^{2+}形成配合物获得与游离 ATP 不同的反应能力，暗示着被生物 体 反 应利

用的可能性是很有意义的。

4.2.5.3 多聚核苷酸

关于多聚核苷酸，已经作了一些富有启发性的理论计算。其中，有关B-DNA分子的四种同聚多核苷酸（聚A、聚G、聚C和聚T）以及考虑到RNA的聚U的能级结构，在考虑了最邻近碱基间的相互作用下已经算出。其办法是，将多聚链作为一维固体，采用通常的固体物理学方法，就容易将所需处理的分子的问题化为求复杂的Hamilton矩阵的特征值问题。矩阵的阶等于分子中π轨道的数目。这些特征值问题还可进一步化为阶数是原来复杂系统阶数两倍的一个对称矩阵的特征值问题。当然，后一步只是一个量子力学的技巧。这里，我们并不讨论具体方法。表4-23列出了由此获得的多聚A的能级结构。

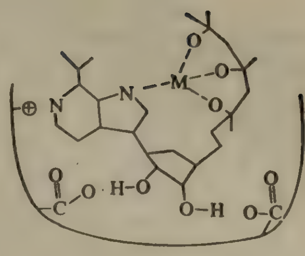


图4-24 ATP—M^{2+}和酶分子的结合模型

表 4-23 多聚A的能级（β单位）

能级		E_{min}	E_{max}	ΔE	E_{MD}
占据能级	1	2.964	2.593	0.37	2.770
	2	2.265	1.993	0.27	2.131
	3	1.805	1.707	0.10	1.754
	4	1.167	1.143	0.03	1.146
	5	1.026	1.000	0.03	1.028
	6	0.576	0.480	0.10	0.530
空能级	I	−0.749	−0.823	0.07	−0.786
	II	−0.984	−1.023	0.04	−1.003
	III	−1.367	−1.006	0.04	−1.386
	IV	−2.008	−2.005	0.02	−2.016

表中E_{min}、E_{max}分别表示能级的下、上限，ΔE是它们的差。E_{MD}是单个腺嘌呤分子用同一参数获得的近似值。占据能级用阿拉伯数字表示，空能级用罗马数字表示。

在堆积碱基之间可分别采用重叠相互作用和Van der Waals型相互作用进行计算。Van der Waals相互作用给出的是稳定能量，但不能给出沿DNA螺旋轴的电子传导的信息。相反，重叠相互作用给出的不是稳定能，但能给出DNA分子中电子传导的信息。不过，两种相互作用的计算结果都能解释减色现象。如果同时考虑两种相互作用，计算结果将使理论图象更符合实际。

近年来，Ladik等[29]对B-DNA结构中所有16种可能堆积形式的堆积能作了计算，他们采用STO-3G基集合实行超分子计算。色散能系用以下London公式确定的：

$$E_{disp}^{AB}=-\frac{3}{2}\sum_{i=1}^{NA}\sum_{j=1}^{NB}\frac{\alpha_i d_j I_i I_j}{(I_i+I_j)R_{ij}^6} \tag{4.11}$$

表4-24列出了B-DNA结构中，16种可能堆积形式中的8种的堆积能和DNA的Z构象

表 4-24　核苷酸碱基间的堆积能(kcal/mol)

	E_{HF}		E_{disp}		E_{tot}	
	B-DNA	Z-DNA	B-DNA	Z-DNA	B-DNA	Z-DNA
G,G*	4.34	1.89	−7.08	−3.82	−2.74	−1.84
G,A	−0.02		−7.22		−7.24	
G,T	2.35		−5.68		−3.33	
G,C	1.07	<u>−2.85**</u>	−5.97	<u>−4.73</u>	−4.90	<u>−7.58</u>
T,G	2.05		−6.31		−4.26	
T,A	1.98	−0.22	−6.69	<u>−2.09</u>	−4.71	<u>−2.31</u>
T,T	4.06		−7.08	−4.16	−3.02	−2.61
T,C	1.66		−7.46		−5.80	

* X，Y表示 $\begin{smallmatrix}Y\\ \uparrow\\ X\end{smallmatrix}$，** 划横线的是Z-DNA中，嘌呤，嘧啶或嘧啶，嘌呤堆积。

中 G,G；G,C；T,A 和T,T 堆积的堆积能的计算结果。

有趣的是，多聚 A最高占据轨道能量的上限和最低空轨道能量的下限小于单个腺嘌呤分子相应的差值。由于被光照引起的电子转移受阻，因此这一事实显不出大的分立光谱实验结果来，但它有助于支持电子的热激发观点。表 4-23 告诉我们，在 B-DNA 中，$\begin{smallmatrix}A\\ |\\ G\end{smallmatrix}$ 堆积具有最负的堆积能，故推测 $\begin{smallmatrix}A\\ |\\ G\end{smallmatrix}$ 堆积最稳定，$\begin{smallmatrix}G\\ |\\ G\end{smallmatrix}$ 堆积在8种堆积中最不稳定。有趣的是在Z-DNA中，$\begin{smallmatrix}C\\ |\\ G\end{smallmatrix}$ 堆积具有最负的堆积能(仅对 4 种堆积情形而言)。理论计算的结果与活细胞中 DNA 碱基堆积对其构象稳定性的影响在实际上符合得如何，无疑还须作更进一步的研究。但是，量子化学计算确实在这类问题的研究中显示出了令人关注的、特殊的作用(参见4.3)。

虽然对于实际存在的 DNA 分子，我们知道的东西并不算多，理论计算又困难重重。但是，我们毕竟在 DNA 分子某些局部性质的理论计算方面取得了进展，至少已经获得了一些定性和半定量的结果。从理论研究的角度，我们可以首先处理一个周期性无限序列 $\begin{array}{c}\text{-A-G-A-G-A-G-}\\ \parallel\ \ |||\ \ \parallel\ \ |||\ \ \parallel\ \ |||\\ \text{-T-C-T-C-T-C-}\end{array}$。这个模型系统向我们表明了碱基间的相互作用，其中既有配对碱基中两个组分间的共平面相互作用，又存在着堆积碱基间的相互作用。如果将堆积碱基之间的相互作用的影响假定为一个平均能级差，取作 0.70 β。由此而得到的一些结果列于表4-25中。

表中，左右两栏是 A-T、G-C 碱基对的 MD 值。第二栏是两对相互作用碱基对的近似值，这些值由于碱基对之间的相互作用有时发生扰动。如果假定能级差为 0.10 β，便可获得 9 个占据能级和 3 个空能级，如表中第三栏所示。

在最高占据能级和最低空能级之间的能量差为：

$\Delta E(-0.803-0.355+0.10)\beta=-1.058\ \beta$

正如处理杂环分子 Hückel 近似下的第一激发态的能量那样，为了与实验数据符合须将 β 积分值取得较大一些，如取作 $\beta=-3.33$ eV。这时 HOMO 和 LUMO 的能量差 $\Delta E=3.52$ eV。这样得到的能量值能够较好地符合实验测定的最大吸收光谱值。由此看来，即使在多核苷酸的

表 4-25　A-T和G-C碱基堆积的MO能量(β)

A-T	A-T / G-C	能　级	G-C
	−1.079	III	−1.079
−0.945	−0.945	II	
	−0.832		−0.832
−0.803	−0.803	I	
	0.355		0.355
0.451	0.451		
0.533	0.533		
	0.554	1	0.554
	0.648		0.648
0.903	0.903		
	0.991		0.991
1.086	1.086	2	
	1.107		1.107
1.149	1.149		
	1.335	3	1.335
	1.730		1.730
1.736	1.736		
1.755	1.755	4	
1.798	1.798		
	2.123		2.123
2.162	2.162	5	
	2.177		2.177
	2.593	6	2.593
2.775	2.775	7	
3.110	3.110	8	
	3.328		3.328
	3.457	9	3.457
3.529	3.529		

情况下，仍然可以进行有效的量子力学计算。

以上对多聚核苷酸的理论计算大多基于 π 电子结构，由此而进行的讨论就不可能是十分令人满意的。近来，Otto，Clementi 等[30,31] 运用 ab initio Hartree-Fock 晶体轨道方法，对四种同聚核苷酸碱基堆积(poly C、poly T、poly A 和 poly G)，两种聚碱基对 poly(A-T)和 poly(G-C)，以及具有腺苷、胸腺苷作为重复单位的两种多聚核苷酸 poly(ASP)和poly(TSP)进行了周期性 B-DNA模型的能带结构计算(有关核酸的能带结构见 4.5)。结果显示：对聚碱基堆积，有重要物理意义的价带和导带非常宽，分别为 0.440～0.789 和0.260～0.820 eV。然而，poly(A-T)和 poly(G-C)，带的样子，位置以及带宽都和生成它们的相应碱基的带有关。对于聚碱基对最小的近似激发能的分析指出，在聚 A-T 对中，HOMO 和 LUMO 能级分别来自 polyA 和

poly T，因此第一跃迁是碱基间型（interbase type）；对于poly(G-C)的分析也得到了类似的结果。

此外，采用Mulliken集居数分析，并按照poly(ASP)和poly(TSP)能级，Otto等计算了从糖磷酸酯单位到腺嘌呤以及到胸腺嘧啶转移的电荷量分别为0.212和0.190 e。这种情形在更大的B-DNA片断(由三个糖单位，两个磷酸根和两个碱基G、C组成)的研究中亦得到。象这种由单位到单位的电荷转移在生物化学过程中担负着重要作用，因此当我们采用某些方法（如分子静电势方法）来研究生物大分子的这种作用时，应当予以注意。

4.3 DNA双螺旋构象的稳定性

现在，我们从上一节对多核苷酸的讨论转入到对DNA双螺旋构象的讨论。

正如我们所反复强调的那样，进行生物分子(最具意义的当然是生物大分子)量子力学计算的最终目的，就是要从计算结果即包含这些生物分子理化信息的分子波函数，推导和解释尽可能多的有关生物分子已知和未知的物理和化学性质，进而推测其生物学功能。

研究提示，造成DNA分子双螺旋稳定构象的主要因素可能是碱基对间的氢键和堆积碱基之间的堆积相互作用与Van der Waals型相互作用。同时，这些相互作用不单决定螺旋构象的稳定性，而且和多聚核苷酸形成螺旋时的专一性以及DNA"转录"为mRNA时的碱基专一性都有着密切关系。因而人们尝试着用分子轨道法和微扰法来处理碱基间的相互作用，如诱导力或色散力、电离势或极化率，重叠与交换积分等，进行高度近似的量子力学计算，以期推测和阐明这些相互作用的本质以及氢键能和堆积能是怎样造成核酸的稳定性和构型的。

首先，作为核酸稳定性尺度之一的实验值是它的熔融温度 T_m。从表4-26可以看到堆积能(S)与氢键能(H)之和与实测的T_m有极好的对应关系。

表 4-26　各种聚合物的碱基间相互作用能和T_m的比较(eV)

聚合物	堆积能(S)	氢键能(H)	(S)+(H)	T_m
(G-C)·(G-C)	−0.201	−0.484	−0.685	95
(G)·(C)	−0.134	−0.484	−0.618	82
(T-G)·(C-A)	−0.163	−0.280	−0.443	73
(T-C)·(G-A)	−0.142	−0.280	−0.422	67
(A)·(T)	−0.131	−0.076	−0.208	48
(A-T)·(A-T)	−0.121	−0.076	−0.197	39

由于碱基组成不同的各种DNA的熵，一般说来具有大致相同的值，因此用T_m中的热焓项来和 $S+H$ 对应是可以说明问题的。由表4-26还可以看出，G-C对和它之外的氢键对的稳定性之差是由H决定的；$G-C$对彼此之间或者在其它氢键对之间的大小关系是由S决定的。有关堆积能的讨论，在4.2.5.2中已经初步接触到了，以下还将继续讨论到它。

关于DNA双螺旋构象问题。首先应当注意的一个事实是：据X射线衍射研究，B-DNA

碱基间的平衡距离为 0.34 nm。那么,相应于这一距离的碱基间的相互作用能是不是最大稳定能呢？量子力学计算结果给出的答案是比较肯定的。从表 4-27 中可以看到,对大部分碱基而言,最大稳定能(能量最负)时算得的平衡距离和实验值 0.34 nm 相近,偏差在 0.05 nm 以内。有趣的是,在 C 上重叠 G 或者 I 时,平衡距离为 0.280 ～0.285 nm,与实验值偏差较大。此外，相同碱基彼此之间重叠时的堆积能都很小,特别是 G、C 的场合成为排斥能[17]。我们的计算也得到类似的结果。这似乎提示我们：在相同碱基彼此重叠的地方很难生成螺旋或者即使形成了螺旋也不稳定。如果事实确系如此，那么又使我们对于建立 RNA 二级结构模型在能量处理上产生疑问[32-34]。真实情况究竟如何？目前还难以下结论。有人认为，相同碱基重叠之处所造成的螺旋的不稳定，也许正是基因中容易发生突变的位点。

表 4-27　碱基重叠的平衡距离和堆积能(eV)*

碱基重叠	平衡距离 (nm)	在平衡距离的堆积能
C, G	0.280	−0.182
C, I	0.285	−0.187
C, A	0.305	−0.087
T, G	0.315	−0.091
G, A	0.325	−0.115
G, C	0.328	−0.147
A, A	0.328	−0.082
G, G	0.352	0.010
C, C	0.368	0.020
T, T	0.368	−0.028

* 参见表4-23。

图 4-25 是 G 上重叠 C 时的能量对距离所作的曲线。根据这些曲线分析,近距离排斥能 E^{12}(二次交换能)、E^{14}(四次交换能)和吸引能 E^{20}(诱导和色散能)以及 E^{10}(静电能)，它们汇合之处,得到了获得最大稳定能时的平衡位置为 0.328 nm。这也说明了 B-DNA 中，堆积碱基间的平衡距离为 0.34 nm 的合理性。在这一距离内,其相互作用使得 DNA 分子构象最稳定,而相应的能量值又是所有相互作用的极小值,这符合能量极小值原理，即:凡是现实发生的都有向能量极小值处靠近的趋势。

B-DNA 双螺旋构象的另一个显著特征是：每个碱基按旋转 36°堆积。10 个碱基组成一个螺距。这样，又给碱基的堆积和几个碱基对堆积之间的关系带来了新的内涵。图 4-26 给出了随相邻碱基间旋转角角度变化的堆积能和聚合物的稳定性。

由图 4-26 可见,A-T共聚物堆积的碱基间的旋转角和天然 B-DNA 相同，为 36°时,能量有极小值,是最稳定的。这就再一次支持了 B-DNA 双螺旋构象是符合能量极小值原理的稳定构象。并且,这种 A-T 共聚物的 X 射线图和 DNA 的图象几乎相同,理论与实验符合得很好。另一方面,A、T 均聚物的螺旋结构和 DNA 的 X 射线分析结果不同,这跟这类均聚物能量计算结果极小值在 54°附近并不矛盾,倒是从另一个侧面验证了能量极小值原理。此外,从图中还可以看

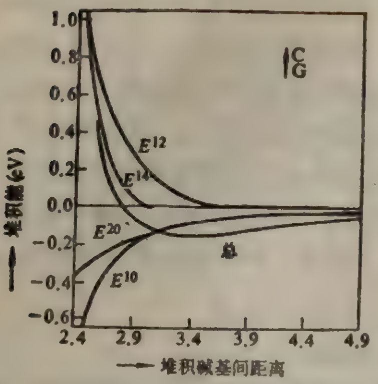


图 4-25 堆积能和碱基间距离

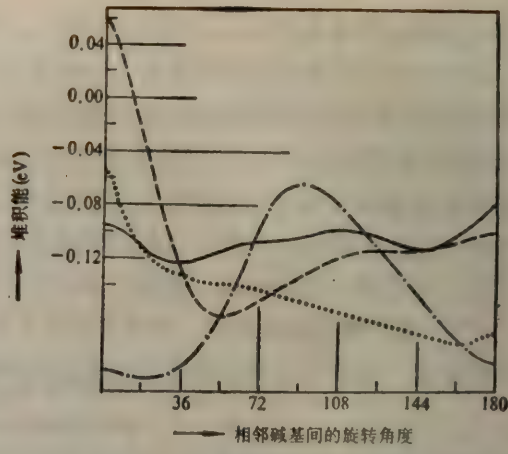


图 4-26 随相邻碱基间旋转角角度变化的堆积能和聚合物的稳定性

——A-T共聚物，——A，T均聚物

-·-·G-C共聚物，······G，C均聚物

到，G、C 均聚物在 36°～72°附近没有极小值，在 160°表现异常值。根据能量极小值原理，这类聚合物将不具有螺旋结构，但事实如何？至少在目前还不能给出确切的解答。

根据以上讨论，我们可以看到，从量子化学对碱基堆积能的计算，在一定程度上很好地说明了 B-DNA 双螺旋构象的稳定性满足堆积能的极小值原理。但是，仅限于对碱基堆积能的计算无疑是远远不够的。其中，有关碱基对间的共平面相互作用以及构象能的计算，读者可参阅本章的后续部分和第十一章“生物分子的空间构象和分子设计”。

4.4 核酸的辐射效应和光生物学问题[10]

生物体受到紫外线照射后会造成各种各样的损伤，如失活、突变、致癌等等。而与此相关的大多数生物效应与 DNA 在分子水平上的物理的、化学的变化有关，由于这样的联系，促使核酸的光生物学飞速发展起来。其中，研究在辐射的影响下由核酸碱基生成的自由基结构与性质以及在光照下胸腺嘧啶二聚体的生成机制已成为重要课题。以下我们仅就在光照下胸腺嘧啶二聚体的生成机制方面的问题，从量子生物学的角度进行分析讨论。

实验发现，经紫外线照射后，DNA 结构中的两个胸腺嘧啶可以结合成二聚体。这种二聚体是正常细胞内 DNA 合成的障碍。如果在细胞内 DNA 复制前这种障碍不被清除，就会产生突变[35]。现在，不仅胸腺嘧啶二聚体的生成已经基本搞清，而且也观察到尿嘧啶、胞嘧啶二聚体的生成。更甚者，是胸腺嘧啶-尿嘧啶、胸腺嘧啶-胞嘧啶、胞嘧啶-尿嘧啶二聚体的生成也有不少研究报导。但是，一般认为嘧啶二聚体中，又以胸腺嘧啶二聚体的生物学作用最为重要[36-39]，故我们将重点放在对胸腺嘧啶二聚体的讨论。

对于胸腺嘧啶二聚体的生成在 280 nm 附近照射效率最高。这种生成物在 240 nm的波长光

的照射下还原为原来的单体：

$$T + T \underset{240\ nm}{\overset{280nm}{\rightleftharpoons}} \widehat{TT}$$

下面几个例子可以很好地说明胸腺嘧啶二聚体的生成与生物活性的关系：

(1) 在 UV 照射下，某些细菌容易死亡，而其他的却有很强的抵抗力。这是因为后者具有所谓“修复”能力，把胸腺嘧啶二聚体从 DNA 上切除后修复，对前者而言，由于没有这种修复能力，因而死亡。

(2) A-T 含量越多，对 UV 的敏感性越高。

(3) *Haemoplilus* 流感转化 DNA 在 280 nm 辐射线照射下失活，而在 240 nm 辐射线照射下活性恢复，这与胸腺嘧啶二聚体有相同的波长依存性。已经知道，在这种情况下，50% 失活是基于二聚体的生成。

(4) 二聚体的生成与光照胸腺嘧啶二聚体产生的不成对电子密度成正比，与基态分子 T 和激发态分子 T* 结合而成的所谓过渡态的离域能成正比。离域能越大，反应活化能越小，就越容易生成二聚体。

(5) 胸腺嘧啶分子成为激发三重态时，5、6位置(二聚体的生成位置，如图 4-27 所示)的 π 键级比之基态最小。因此，胸腺嘧啶二聚体大部分是经过激发三重态而不是激发单线态生成。

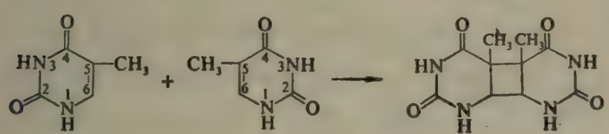


图 4-27

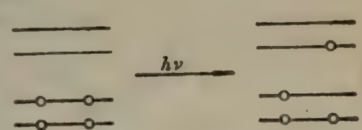


图 4-28 光照射下电子的激发

Pullman 等认为[40]，由于光照，胸腺嘧啶被激发，其分子内电子分布变成容易生成二聚体的状态。如图 4-28 所示，由于光照射激发一个电子，生成两个不成对电子。由于不成对电子的分布，5、6 键的值非常大，因而容易形成二聚体(见表 4-28)[17]。

对胸腺嘧啶二聚体的生成，就量子生物学的研究看，Pullman、Setlow、福井和永田等许多人已经做了很好的工作。例如，Pullman 等用简单分子轨道法(HMO 法)首先研究了第一激发态未偶合的电子分布(图4-29)。

从图 4-29 可见，鉴于不成对电子的分布，键 5、6 的值非常大。可以认为这些不成对的电子参与二聚体的生成。实际上，二聚体的生成能与不成对的电子密度有很好的对应关系(表4-29)。二聚化的产率与C 5—C 6 键电子总浓度平行。唯一的例外是 6-氮胸腺嘧啶。如此，我们可以把不成对的电子密度作为激发状态反应性的尺度。

其次，作为激发态的第二个特征是C 5—C 6 键上键级的减小并且二聚化与这种减小近乎直接相关(见表4-29)。

由上述讨论可以看出，在这些嘧啶之中电子的激发看来是局限在C 5—C 6 双键，并且激发的特点明显地与光二聚化作用的速度和机制有关。

表 4-28　激发状态下两个不成对电子的5、6键位置的分布和二聚体生成能的比较

二聚体生成能	化　合　物	5、6位置的两个不成对电子分布之和*
大	胸腺嘧啶	1.21
	尿嘧啶	1.25
	6-甲基尿嘧啶	1.21
	乳清酸	1.12
中	N1，N2二甲基尿嘧啶	1.16
	异胞嘧啶	1.16
	5-氨基尿嘧啶	1.05
小	5-甲基胞嘧啶	0.83
	胞嘧啶	0.86
无	2-巯基胸腺嘧啶	0.71
	6-硝基尿嘧啶	0.64
	6-氮胸腺嘧啶	1.14

* 这是阳离子基团，阴离子基团自旋密度不成对电子之和。

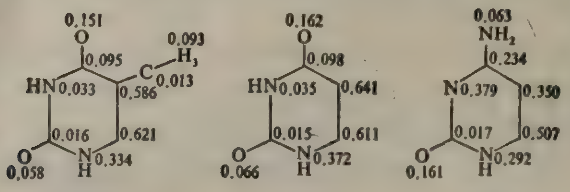


图 4-29　第一激发态未偶合的电子分布

表 4-29　C5—C6 键级

化　合　物	基　　态	第一激发态	差
胸腺嘧啶	0.811	0.287	0.524
尿嘧啶	0.819	0.298	0.521
S-甲基尿嘧啶	0.790	0.296	0.494
乳清酸	0.809	0.323	0.486
N1,N2-二甲基尿嘧啶	0.811	0.348	0.463
异胞嘧啶	0.817	0.319	0.498
5-氨基尿嘧啶	0.806	0.300	0.506
5-甲基胞嘧啶	0.749	0.430	0.319
胞嘧啶	0.758	0.459	0.299
2-巯基胸腺嘧啶	0.804	0.474	0.330
6-硝基尿嘧啶	0.733	0.578	0.155
6-氮胸腺嘧啶	0.831	0.295	0.536

进而永田等根据福井等提出的芳香族化合物激发状态的反应性理论，具体地处理了胸腺嘧啶二聚体生成并提出了如下的生成机制[41,42]：

$$A + A^* \xrightarrow[a]{(A\cdots A^*)} (AA)^* \xrightarrow[b]{} \widehat{AA}$$

即基态分子A和激发态分子A*结合，经过所谓过渡态(A···A*)产生二聚体 $\widehat{AA}$ ，a 是限制反应速度的阶段。可以预料在和基态相同的情况下，过渡态的离域能越大则活化能越小，则容易引起反应。这种情况的离域能如图 4-30 所示，并根据微扰理论按下式计算了胸腺嘧啶和其他一些化合物的离域能：

$$\delta E = \sqrt{(C_r^{o2} d_s^{lu})^2 + (C_r^{o1} d_s^{ho})^2 \nu} + \left\{ \sum_j^{\text{未占}} \sum_i^{\text{全体}} \nu_i - \sum_j^{\text{占有}} \sum_i^{\text{全体}} (2-\nu_i) \right\} \frac{(C_r^i C_s^j)^2}{2(\varepsilon_i - \varepsilon_j)} \nu^2 \quad (4.12)$$

式中，C_r^i是第 i 分子轨道的第 r 原子轨道的系数，ε_i，ν_i是 i 分子轨道的能量和 π 电子数。ν 是二聚体生成反应位置的共振积分。在计算中已略去 ν 的二次项。

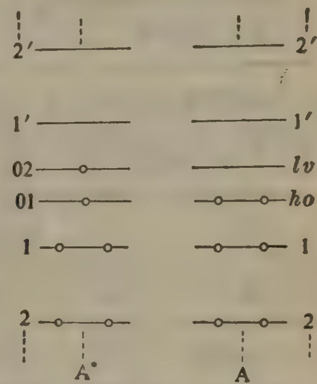


图 4-30 基态分子 A，激发分子 A* 的分子轨道

计算结果与实验值基本符合(表4-30)。从表中可以看出，除萘外，离域能和二聚体生成能对应得很好，表明激发态的反应也和基态一样，底物和试剂间(此时基态、激发态两分子之间)的电子流动是稳定化的重要因子。有趣的是胸腺嘧啶、尿嘧啶等核酸碱基的二聚体生成可以从 δE 值来和其它共轭化合物的二聚体相比较。与容易起光二聚化反应的化合物蒽比较，胸腺嘧啶或尿嘧啶的反应性更大。

除以上列举的工作外，还可以列出其它一些研究工作，尤其是运用较精确的分子轨道法所作的工作。但是，关于胸腺嘧啶二聚体的生成机制，迄今仍有许多不清楚的地方。例如，胸腺嘧啶二聚体的生成到底是基于怎样的力？联结在C5位的甲基，在生成二聚体时其空间位阻和电子效应如何？等等，还有待作出量子生物学研究的回答。当然，核酸的辐射效应和光生物学问题的研究就更加复杂和需要更长的时间了。

4.5 核酸的能带结构[10]

核酸与蛋白质分子一样也具有能带结构。自 60 年代以来，曾用 CNDO/2 法、MINDO/2 法和 P-P-P 法等分子轨道法计算过 DNA 的能带结构。

有实验报导，DNA具有半导性[43]。这可能是由于DNA分子中的共轭结构和大量相互作用的非偶电子云，为半导性的出现提供了条件的缘故，这也正是一般的有机半导体的特征。非偶电子的能级相互影响形成宽阔的能带，并且扩展到整个大分子的范围。激发这些非偶电子所需克服的能垒(激发能)较小，因此非偶电子容易被激发。激发后，就在整个大分子上形成一个传导带，这就出现了半导性。

表 4-30 各种共轭化合物的生成能和离域能的比较

化 合 物	二聚体形成的位置	离 域 能（γ单位）	二聚体生成能*
尿嘧啶	{ 4-4 5-5	0.761	+
胸腺嘧啶	{ 4-4 5-5	0.701	+
苊烯	{ 1-1 2-2	0.568	+
蒽	{ 9-9 10-10	0.547	+
2-甲基蒽	{ 9-9 10-10	0.544	+
1,2-二苯乙烯	{ 5-5 6-6	0.543	+
9-甲基蒽	{ 9-9 10-10	0.541	+
9,10-二甲基蒽	{ 9-9 10-10	0.534	+
萘	{ 1-1 4-4	0.512	−
苊	{ 1-1 8-8	0.483	+
四氮烯	{ 5-5 12-12	0.417	±
萘醌	{ 2-2 3-3	0.414	±
胞嘧啶	{ 4-4 5-5	0.406	±
戊省	{ 6-6 13-13	0.399	±
芘	{ 3-3 8-8	0.384	−

* +，−表示二聚体形成的有无，±表示弱活性。萘为−，但已知能形成激发态二聚体(excimer)。

Ladik 等[44,45]用 HMO-LCAO 方法对 DNA 中的能带结构进行了计算。结果表明，DNA 中的能带结构是以(A-T)和(G-C)碱基对的周期交互堆积而成，价带和导带（最低空带）之间的禁带宽度是

$$E_G = 3.52\ \mathrm{eV} \text{ 或 } 2.69\ \mathrm{eV}$$

这取决于碱基间氢键上的氢原子是否定位于碱上面而不同。Ladik 等认为，DNA 分子中的电子也可通过由杂质或结构缺陷而产生的中间能级达到导带。

Suhai用CNDO/2 CO 法计算了 DNA 糖-磷酸骨架并比较了糖-磷酸主链(聚 SP)与同聚核苷酸(polyA、polyT 和 polyC)价带和导带的位置及宽度，结果见表 4-31。其中 ε_{min} 表示带的下限，ε_{max} 表示带的上限，$\delta\varepsilon$ 表示带的宽度。从表 4-31 所列数据可以看到，poly SP 的价带和导带

表 4-31　聚A、聚T、聚C和聚SP的价带与导带

体　系	价　带 (eV)			导　带 (eV)		
	e_{min}	e_{max}	δe	e_{min}	e_{max}	δe
poly A	−9.749	−9.575	0.174	3.712	3.951	0.239
poly T	−10.160	−9.642	0.518	1.810	2.022	0.211
poly C	−10.692	−10.527	0.165	2.875	2.966	0.09
poly SP	−12.644	−12.591	0.053	−1.240	−1.099	0.141

宽度(分别为 0.053 eV和 0.141 eV)并不比同聚核苷酸的小多少。因此，可以设想它们和同聚核苷酸有相近的导电性质。

此外，Suhai 等还计算了由 PO_4' 部分和 K^+离子形成双层引起的不均匀电场对核苷酸能带结构的影响。表 4-32 给出了在有无电双层时 CNDO/2 和 MINDO/2* 法计算所得到的四种多核苷酸导带和价带的宽度。

表 4-32　四种多核苷酸的导带和价带 (eV)*

体　系	δe_v	δe_c	ΔE	δe_v	δe_c	ΔE
	(无电场)			(有电场)		
poly A	0.26	0.11	13.97	0.28	0.36	13.83
poly T	0.05	0.13	14.19	0.11	0.44	14.00
poly G	0.26	0.01	12.19	0.32	0.43	11.94
poly C	0.06	0.23	13.26	0.10	0.75	12.75

* δe_v(价带宽)，δe_c(导带宽)，ΔE(价带与导带之间的距离)

从表中可以看到，导带和价带的宽度增大了。由此可见，PO_4'-K^+ 离子电双层对堆积碱基对的带结构有重要作用，所以在处理 DNA 电子结构时必须考虑这一点。

用整个核苷酸作为重复单元的能带结构计算给出了从糖到核苷酸碱基有～0.2 e 的电荷转移，转移的电荷量与核苷酸的碱基类型无关。研究表明，多核苷酸的价带(n^*)和导带(n^*+1)是“碱基堆积”型，而n^*-1 带和n^*+2 带是“糖-磷酸根链”型。就这些情况而言，简单的HOMO(给予体)⟶LUMO(接受体)的电荷转移图象并不成立。通过对 n^*+1 和 n^*-1(即最高充满糖-磷酸根型带) 能量间隔的计算。

Ladik、Otto 等指出，由于较大的能量间隔 (大于 13 eV)，在 Hartree-Fock 水平上能够排除 B-DNA 中生成自由电荷载体的可能性。因此，如果堆积碱基是周期性的，B-DNA 将成为固有半导体[29]。

* 表 4-31 为 CNDO/2 CO 法计算所得数据，用 MINDO/2 法计算结果类似。

4.6　核酸和蛋白质的相互作用

在生物系统中，核酸和蛋白质的相互作用是最重要的相互作用，它决定和联系着多种重要的反应，如在遗传信息传递过程中的核酸和蛋白质的相互作用。以下我们将分作几个问题来叙述这种相互作用的量子生物学研究。

4.6.1　碱基对与丝氨酸的专一识别

这方面，Hosur 等[46]在进行了聚L-丝氨酸与聚 U、聚 A 以及聚 A-聚 U 复合物键合的NMR和 UV 实验之后，对聚 L-丝氨酸与 A-U、G-C 碱基对相互作用作了理论计算和讨论。阐明了聚L-丝氨酸与碱基对结合的专一识别中蛋白质骨架的作用。从计算结果看，他们对多肽骨架和核酸碱基间的相互作用模型(图 4-31)作了详尽的分析，强调了碱基序列专一识别的重要性。定性的结论是：相互作用模式对专一识别的贡献取决于这些模式是否能在任一序列中识别 A-U(T)，U(T)-A，G-C 和 C-G 碱基对。考查这些碱基对和多肽骨架之间在大沟中的氢键(图 4-32)显示出：A-U(T)和 C-G 碱基对具有相同的氢键方向(N→O)，而 A-U(T)和 G-C 则显示出相反的氢键方向。这说明在成键位置具有固定几何模型的蛋白质肽骨架由于不同的氢键方向使之能容易地根据 G-C 辨别 A-U(T)或根据 C-G 辨别 U(T)-A。

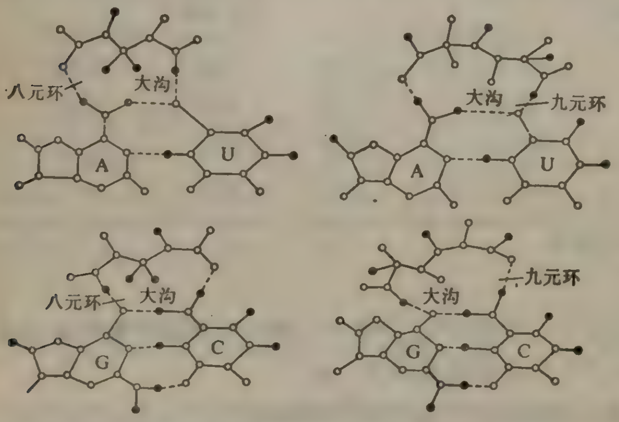


图 4-31　肽链通过 H-键桥跨越至碱基对，并产生两种不同的图形——八元环和九元环。

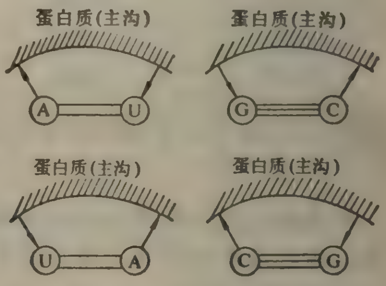


图 4-32

表 4-32 给出了对不同的空间允许的多肽骨架构象，肽和碱基对之间的相互作用能。必须注意，关于相互作用能的这些结果反映的是在真空中的情形。而在水溶液中肽与碱基对结合会导致水的组织的改变，从而引起体系溶剂化能、熵和构型熵的改变。遗憾的是，这些贡献目前还未给予充分的考虑。从表 4-33 可以看出，A-U 和 G-C 对具有极不同的相互作用能。这些结果将有助于聚 L-丝氨酸结合碱基对的大多数实验结果的解释。现有的实验表明，聚 L-丝氨酸跟聚A、聚 U 和聚 A-聚 U 结合是通过涉及肽链和核苷酸(或碱基对)的 H 键实现的。

表 4-33 对不同的空间允许的肽骨架构象，肽和碱基对间的相互作用能(kcal/mol)

碱基对	环大小	二面角①				相互作用能	构象能	稳定化能	氢键键长(nm)	氢键键角(度)
		ϕ_1	ψ_1	ϕ_2	ψ_2					
A-U	8	200	330	300	0	−6.7	6	−0.7	0.303	13.8
									0.322	8.4
G-C	8	210	320	90	130	−13.2	7	−6.2	0.306	3.0
									0.295	0.3
A-U	9	80	150	50	270	−8.5	2	−6.5	0.285	13.4
									0.286	3.2
A-U	9	70	210	340	270	−6.7	4	−2.7	0.287	6.7
									0.287	0.7
A-U	9	60	230	330	270	−6.1	6	−0.1	0.287	6.7
									0.287	1.6
G-C	9	20	200	0	330	−14.1	4	−10.1	0.288	16.3
									0.286	35.4
G-C	9	60	70	180	270	−9.5	4	−5.0	0.293	6.0
									0.285	3.5

① ϕ和ψ角定义当取反式位置时作为零：$\psi = N—C^{\alpha}—C'—N$和$\phi = C'—N—C^{\alpha}—C'$。

不难看出，肽链跟 G-C 对的相互作用强于 A-U 对的相互作用，这是由于 A-U 对和 G-C 碱基对电荷分布不同所致。

4.6.2 B-DNA 与多肽的相互作用

为了研究与 DNA 结合的蛋白质骨架的成键性质，Otto、Clementi 等[47]参照由实验得到的结构数据，建立了一种 B-DNA 与多肽相互作用的模型。

在处理时，对于周期性单股 DNA 螺旋仍用聚 C、聚 T、聚 A、聚 G 和聚(Asp)和聚(Tsp)表示，而 Watson-Crick 碱基对 G-C 和 A-T 以及二核苷酸 Asp-Tsp 则选择来作为 DNA 双螺旋模型大分子的基本单位。对于多肽则选取聚甘氨酸并采用六种构象，它们是：充分伸展的 β-折迭形式的聚甘氨酸(PGFE)，α-螺旋(PGαH)，以及螺旋轴与 DNA 一致，螺旋半径约为 1.5 nm，螺距分别为 B-DNA 螺距的三倍、两倍、等长和四分之一的聚甘氨酸螺旋(记作 PG 101，PG 67，PG 33，PG8)。

首先，对聚甘氨酸的上述六种构象进行能带结构计算，根据每单位(含二个甘氨酸)的总能量，推测出稳定性由大到小的次序是 PGFE，PGαH，PG 33，PG 8，PG 101。

依据分子间相互作用的一般理论，运用相互自洽场(the mutually consistent field)方法[48]和微扰理论表达，对 B-DNA 多肽相互作用模型的理论计算结果归纳在表 4-34 和 4-35。表中第二列给出静电、极化、交换、电荷转移四种能量，第三列为色散能，最后一列为相互作用能的总和。

分析理论计算所获得的结果，再根据有关核酸组蛋白复合物的分子亚单位计算研究以及阻遏物蛋白的实验迹象，Otto 等推断出适合于实际体系的一般结论。

表 4-34 充分伸展构象的聚甘氨酸和周期性单、双股DNA螺旋间相互作用能贡献（kcal/mol）

B-DNA 模 型	充分伸展的聚甘氨酸 (PGFE)		
	$E_{el+pol+ex+CT}$	E_{dis}	E
poly A	−2.9	−2.4	−5.3
poly G	−3.1	−2.0	−5.1
poly T	−2.7	−4.4	−8.1
poly C	−3.6	−3.9	−7.5
poly (ASP)	−24.6	−1.7	−26.3
poly (TSP)	−21.0	−1.7	−21.7
poly (A—T)	−9.6	−6.9	−16.5
poly (G—C)	−7.7	−5.8	−13.5
poly (ASP—TSP)	−24.4	−2.8	−27.2

表 4-35 B-DNA单、双股螺旋，聚(ASP)和聚(ASP-TSP)分别与不同螺旋构象的聚甘氨酸的相互作用能（kcal/mol）

聚甘氨酸螺旋	B-DNA模型					
	poly (ASP)			poly(ASP-TSP)		
	$E_{el+pol+ex+CT}$	E_{dis}	E	$E_{el+pol+ex+CT}$	E_{dis}	E
PGαH	−14.4	−1.1	−15.5			
PG101	−4.8	−2.8	−7.6	−0.8	−5.1	−5.9
PG67	−10.7	−1.7	−12.4	−21.9	3.3	−18.6
PG33	−55.3	−0.9	−56.2	−27.6	−4.0	−31.6
PG8	−5.6	−3.2	−8.8	−2.4	−6.1	−8.5

(1) 充分伸展的β多肽骨架是最稳定的，且按照相互作用模型，每隔一个就有另一个氨基酸残基指向DNA螺旋并将与磷酸根的邻接碱基相互作用。不过这种可能性相当小，因为PGFE与DNA仅在两链准确配合在一起时才形成强有力的定域键，而一旦这种情形发生，属于充分伸展构象并且有中性或亲水性侧链的肽单位将会有关于转动的高度灵活性。显然维持蛋白质功能的结构不能归因于这些松弛相联的区域和可流动的多肽序列。

(2) 对α-螺旋构象，每个DNA螺距的独自相互作用能仍小于PGFE。呈这种构象形式的一个较长的多肽序列相对于DNA而言会使这一区域变得更易为化学反应攻击。此外，每隔4/5个氨基酸侧链就与DNA亚单位的合适的成键位置相冲突。于是较好的成键性质可以通过导入一种形变为线性多肽链而获得，结果α-螺旋服从于某种程度的DNA螺旋而伸展。

(3) PG 33与DNA间的最大相互作用能有力地支持了DNA能够在多肽链的构象上引入它的螺旋对称性的假设。这一假设是由肽和核苷酸碱基间存在的化学计量法提议的，符合于在合成的基本同聚肽-DNA复合物中所得到的1:1关系。当然，由于PG 33的半径大于DNA螺旋的半径，为构成PG 33螺旋，比核苷酸单位多两个甘氨酸是必要的。其它多肽螺距不同于B-

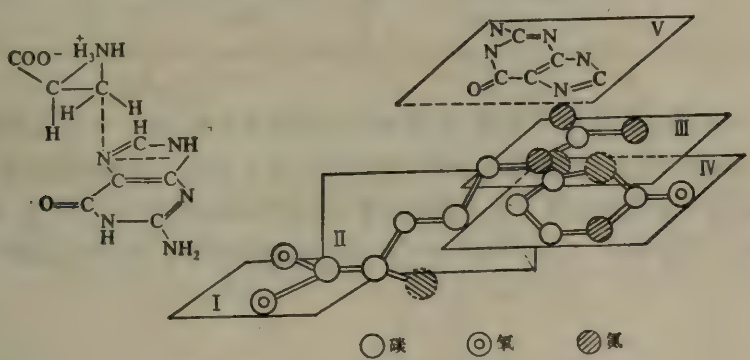
DNA 参数的情形，其可能构型的稳定性来自侧链与 DNA 的相互作用，因而主要包含甘氨酸和丙氨酸的序列将再次成为不稳定的复合物区域。PG 33 则不同，即便缺乏合适的氨基酸残基，多肽骨架对整个复合物的稳定化仍能作出最基本的贡献，而两链亚基间的相互作用则是使键强度增加。然而，必须指出，上述三点结论仍旧存在着明显的不足之处。究其原因，除其它因素外，就理论计算本身而言，是由于作了几点近似，即某些亚基间可能存在的相互作用的忽略，多肽骨架完全周期性的假定以及周围水合离子效应的忽略。

4.6.3 遗传密码子与氨基酸的相互作用

在本书中，有关遗传密码子的叙述可参见 7.2.1.1、7.2.1.3 和 7.2.2.2、7.2.2.3。这里，我们仅就遗传密码子与氨基酸相互作用的理论计算作一简述。

首先，需要指出的是，我国科学工作者乐树云、江寿平已经对密码子与反密码子的识别和相互作用进行了理论计算[49]。

作者等在如图 4-33 所示的模型的基础上，根据(4.13)式作了计算[8,50]。



A *B*

图 4-33 密码子-氨基酸相互作用模型

(*A*) 丙氨酸-密码子，图中仅示出第一碱基 G。

(*B*) 精氨酸-密码子：I. 羧基 α-碳平面，II. 碳链平面，III. 胍基平面，IV. 第一碱基 C 平面，V. 第二碱基 G 平面。

$$E_M = E_{\rho\rho} + E_{\rho\alpha} + E_L \tag{4.13}$$

式中，$E_{\rho\rho}$ 表示单极子-单极子相互作用；$E_{\rho\alpha}$ 表示单极子-诱导偶极子相互作用；E_L 表示色散相互作用能。它们分别为：

$$E_{\rho\rho} = \sum_{i\in A}\sum_{j\in B}\frac{\rho_i\rho_j}{R_{ij}} \tag{4.14}$$

$$\begin{aligned}E_{\rho\alpha} &= -\frac{1}{2}[\boldsymbol{\mu}_{A\to B}\boldsymbol{E}_{A\to B} + \boldsymbol{\mu}_{B\to A}\boldsymbol{E}_{B\to A}]\\ &= -\frac{1}{2}[\alpha_B(\boldsymbol{E}_{A\to B})^2 + \alpha_A(\boldsymbol{E}_{B\to A})^2]\end{aligned} \tag{4.15}$$

因为 $\boldsymbol{\mu}_{A\to B} = \alpha_B\boldsymbol{E}_{A\to B}, \boldsymbol{\mu}_{B\to A} = \alpha_A\boldsymbol{E}_{B\to A}$

且

$$E_{A\to B}=\sum_{i\in A}\frac{\rho_i}{(R_{iB})^3}\boldsymbol{R}_{iB},$$

$$E_{B\to A}=\sum_{j\in B}\frac{\rho_j}{(R_{jA})^3}\boldsymbol{R}_{jA}$$

$$E_L=-\frac{3}{2}\frac{I_A+I_B}{I_A+E_B}\frac{\alpha_A\alpha_B}{R^6_{AB}} \tag{4.16}$$

结果如表 4-36 所示。

表 4-36 密码子-氨基酸相互作用能相对值比较(kcal/mol)

密码子	精氨酸	密码子	丙氨酸
CGU	−3.837	GCA	−6.889
CGC	−1.492	GCG	−8.392
CGG	−8.436	GCC	−3.912
CGA	−5.479	GCU	−3.190

分析表 4-36 中的能量值可以看到,虽然遗传密码是简并的,即一个氨基酸能被多个密码子所限定,但其中可能存在着一个最优势密码,它的转译能力最强。譬如由确定精氨酸的六个简并密码中所计算的四个密码子来看,CGG 最占优势,但是它的反密码子 GCC 在确定丙氨酸的四个简并密码中却不占优势(−3.912 kcal/mol),在丙氨酸密码中,GCG 占优势(−8.392 kcal/mol)。

4.7 核酸和水的相互作用

在量子生物学研究领域,人们日益重视溶液介质对生物大分子结构与功能的影响。毫无疑问,最有意义和最有价值的理论计算,是对最接近活细胞中实际存在的大分子结构和反应活性部位直至这些大分子与其它分子相互作用的计算。诚然,这样的计算在目前尚存在着某些实际上的不可行性。但是整个量子生物学的研究,从认识论到方法学都在朝着解决这些问题的方向发展。

像核酸和蛋白质这样一些生物大分子在活细胞中的存在形式,过去给人的印象似乎总是相当坚固的。这也并非没有一点道理,因为磷酸基团对核苷酸的坚固性是有力的说明,而且任何功能运动都可以通过围绕磷酸二酯键的旋转而产生。因而根据单核苷酸是坚固的见解,就得出了核酸的功能单位——多核苷酸也是刚性的和坚固的结论。可是,实际存在的核酸或蛋白质总是处于一定的溶液介质之中。事实上,来自各方面的证据都指出处于真实溶液状态下的核酸和蛋白质是柔性的,即所谓柔性分子。因此,研究核酸和水的相互作用就成了量子生物学的重要课题,尽管这样的研究在电子水平尚未取得令人鼓舞的成果,但进展却是迅速的。

例如，Pullman 等[51]通过量子力学计算获得了核酸中单个碱基同水相互作用产生的围绕碱基周围的第一(也是唯一)水合层(first hydration shell)的最稳定状态，如图 4-34 所示的水合状态图。

在仔细地讨论这些水合层的结构之前，我们不妨先来看一看水本身相互作用的情形。在理想的 H_2O-H_2O二聚体中。水分子间的相互作用能是－6.7 kcal/mol，而在中心水分子周围有 4 个水分子包围的情况下，水分子间的相互作用能是－5.1 kcal/mol。我们可以分别用这两个能量值作为区别水分子是否和“碱基”结合的界限。从图 4-34 可以看到，其中的相互作用能都大于－5.1 kcal/mol，因此这些水分子都属于碱基水合层之内，而用－6.7 kcal/mol作为界限，一些小于这一数值的水分子就将不属于碱基的水合层之内。

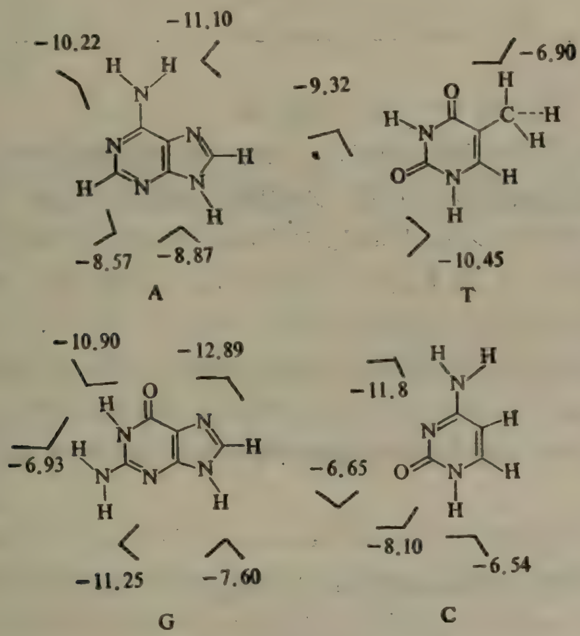


图 4-34 核酸碱基的水合状态

Pullman 同时采用了－5.1 和－6.7 kcal/mol 这两个数值来作为“结合”水分子的最低下限，获得了大致相同的结果。不同之点也只是由于计算标准的选取不一样引起的。因此，我们只介绍按第一个标准获得的结果。有关碱基-水相互作用的必要参数收集在表 4-37 中。它们是水合层内的水分子数、碱基-水相互作用能以及为建立碱基-水相互作用而断开水-水相互作用所需要的水合能。

由表 4-36 可见，A-T 和 G-C 碱基对与它们的二聚体 A＋T 和 G＋C 相比，水合能分别减少了 5.8 kcal/mol 和 18.8 kcal/mol，与水合层内减少的水分子数成对应关系。减少的水分子数分别为一个和四个。此外，A-T 碱基对较 G-C 碱基对在水合层内多含一个水分子。

江寿平等[52]利用分子间相互作用的势能函数研究了核酸碱基 C、A 与水的相互作用。结果表明，碱基C、A与水相互作用能的极小值及其最优的水合作用位置(图 4-35 和 4-36)与Pullman

表 4-37 碱基和碱基对水合状态能(kcal/mol)

碱 基	水合层内水分子数	与水相互作用能	水 合 能
A	4	−38.8	−18.4
T	3	−26.7	−11.4
G	5	−49.5	−24.1
C	4	−33.1	−12.7
A-T	6	−54.6	−24.0
G-C	5	−43.5	−18.0
A+T	7	−65.5	−29.8
G+C	9	−82.6	−36.8
(A+T)-(A-T)			−5.8
(G+C)-(G-C)			−18.8

等人所获得的结果接近，但在方法上作了很大简化(参见 9.2.2.1)。

研究表明，为了使形成的互补碱基对总体运动更趋平衡，也必须考虑由于碱基间氢键引起的相互作用能。Pullman 等[51]对此也作了计算，其结果是：碱基内相互作用能 A-T 碱基对为 −14.6 kcal/mol，G-C 对为 −22.9 kcal/mol。这个数值与从质谱仪测定的真空状态下碱基结合能(A-T 对是 −13.0 kcal/mol，G-C 对是 −21 kcal/mol)的实验值十分接近[53]。由此可以概括为：

(1) 水合 G-C 对较水合 A-T 对更稳定，它们的总相互作用能（碱基内+水合）分别为 −40.9 和 −38.6 kcal/mol。

(2) 水合 A-T 对和 G-C 对相对于它们的共聚物(A+T)和(G+C)更稳定，稳定能分别是 −8.8 和 −4.1 kcal/mol。

从这里可以看到，水合作用使共聚物相对于碱基对丧失了部分稳定性，对于 G-C 对更是如此。不过碱基对之间的氢键和堆积碱基间的相互作用使水溶液中的核酸仍然保持构象的稳定。其实，实际状态的核酸并不象我们想象的那么坚固，而是由于水合作用而具有一定的柔性。

无疑，有关溶剂效应的理论计算对于核酸与水相互作用的研究是重要的。过去在处理溶剂效应时，是采用所谓“连续”模型的方法，而用量子力学方法处理是“分离”的所谓“超分子”模型处理，即把溶质和邻近同它相互作用的溶剂分子作为一个整体进行处理。这种方法是由 Pullman 提出的。运用“超分子模型”法可处理溶剂分子，如水在生物分子上最可能的结合位置，从而帮助我们了解溶剂对生物分子的影响。研究表明，腺嘌呤和鸟嘌呤最适的水合位置有二个，而胞嘧啶和胸腺嘧啶有三个(图 4-35、图 4-36)

图 4-36 表示胞嘧啶的水合情况。从图中可见，在胞嘧啶中 N 3 位提供了最有利的水合位置。

江寿平等的计算还提示[52]：核酸碱基(如胞嘧啶和腺嘌呤)的水合作用最优位置和水合作用能的分布是与碱基分子中各原子的净电荷分布紧密相连。在胞嘧啶中因 N 3 原子的净电荷最负(−0.663)而胞嘧啶水合作用的第一最佳位置就是在 N 3 附近的最小值位置 I。在腺嘌呤中带负电荷的各原子的净电荷按其绝对值大小排列的顺序是：N 7(−0.557)，N 1(−0.518)，N3

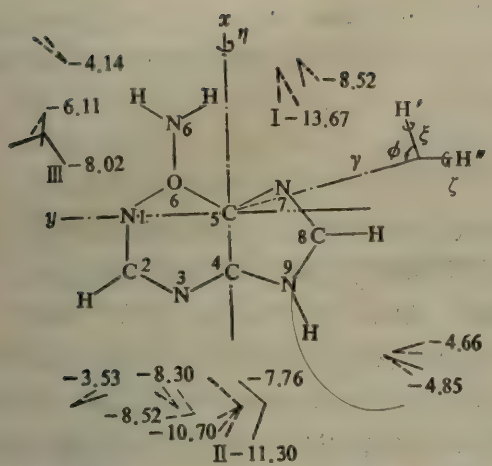


图 4-35 改进势能法计算的腺嘌呤的水合作用图中粗线表示其最佳的水合位置，实线表示与腺嘌呤共面，虚线表示不与腺嘌呤共面

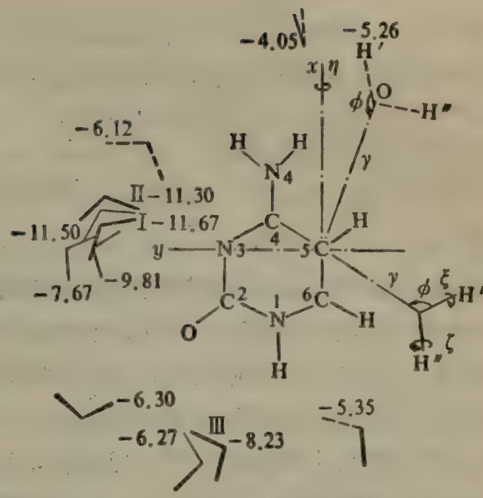


图 4-36 改进势能法计算的胞嘧啶水合作用图中实线表示与胞嘧啶共面，虚线表示不与胞嘧啶共面，粗实线表示其最佳的水合作用位置，能量单位：kcal/mol（详细数据见[52]）。

(−0.510)，N6(−0.439)，N9(−0.192)。而由改进势能法计算的腺嘌呤水合作用的最佳位置的排列顺序是与之对应的，即 N7−C−N6，N3−N9H，N1区域(参见图 4-35 和图 4-36)。由此可见，由改进势能法计算的碱基的水合作用能是与它们的原子净电荷分布紧密相连的。因此可以期望利用现有非经验 ab initio 或 CNDO等方法计算的生物分子的电子结构资料而进行生物分子相互作用的研究。

4.8 核酸的分子静电势研究

在这一节里，我们将转而讨论运用分子静电势(参见 2.2.6)来研究核酸分子的静电整体特征，进而帮助读者理解分子整体结构对其反应位置和反应活性的影响。

我们知道，在分子间的相互作用中，静电相互作用是最基本、最重要的相互作用。由此可见，对静电因素的充分了解无疑将会给我们提供关于生物分子的许多性质。这或许就是分子静电势这一量子化学概念和方法近几年来在国际上日益受到量子生物学普遍重视的一个重要原因。

以下我们将从简单的组成单元逐步过渡到复杂的生物分子本身这一研究程序来予以叙述，

4.8.1 核酸基本组成单元的分子静电势

在具体计算核酸基本组成单元(碱基、糖和磷酸基)的分子静电势时，可根据我们的计算目的和已经具备的计算条件，按第二章中有关分子静电势的计算原理和方法分别采用 CNDO/2 法或从头计算法来获得分子波函数，进而计算分子静电势。

图 4-16 是用等势线表示的碱基平面内的计算结果。

在 4.2.3 中我们已经简介了核酸碱基的分子静电势，对于磷酸二酯单 元和糖等基本单元也可按同样的 ab initio 法或 CNDO/2 法等进 行 处理，并 可把计算扩 展到 各种构 象和稀有碱基[54-60]。

4.8.2 较高单元的分子静电势

作为第二步，是计算由上述基本单元复合而成 的核苷等核酸的高一级单元的分子静电势。

但是，正如第二章中已经讨论过的那样，要想对整个核酸分子 进行严格 的 CNDO/2 或 ab initio 计算，至少在目前是不可能的，或者是极端困难的。近几年来，Pullman 等[60-62]发展了一种近似方法，即用各个个别小单元的势能的迭加来构成整个 系统的 势的方法。对离原点距离不太近的点，采用电子密度的重叠多级展开(overlap multipole expansion,OMTP)方法，可以大大简化计算，降低计算机费用[63]。理论研究和计算实践均表明，OMTP 可以相当满 意 地计算出有关组分单元的任何原子在 0.2 nm 以外的势能的精确值。当然，在 0.2 nm 以内的各点，视情况其势能一般可由 ab initio 法直接得到。

OMTP 方法的要点是用多中心多级展开来代替严格 的 电子 密度分布函数(由 ab initio 法所得)，对任一由 n 个原子组成的单元，展开点为 $n(n+1)/2$ 个。最 简单 的例子就是如图 4-37 所示的 A—T 和 G—C 在其碱基平面上的分子静电势。不难看出，碱基对与 在 单个碱基中的势能极小值的位置改变非常小，它们的吸引力最强的位置顺序是：

$$\text{G-C 对中：} N7(G) > N3(G) \simeq O6(G) > O2(C)$$

$$\text{A-T 对中：} N3(A) > N7(A) > O2(T) > O4(T)$$

其中，鸟嘌呤 N7 位置的负电势最大，其吸引力最强(指对于亲电试剂的进攻)。在形成氢键 的碱基对所在区域中(胞嘧啶的 N3，腺嘌呤 N1，鸟嘌呤的 O6 及胸腺嘧啶的 O4)，孤立 碱 基中的负电势基本上(或几乎)被对面位置的强排斥势抵消了。

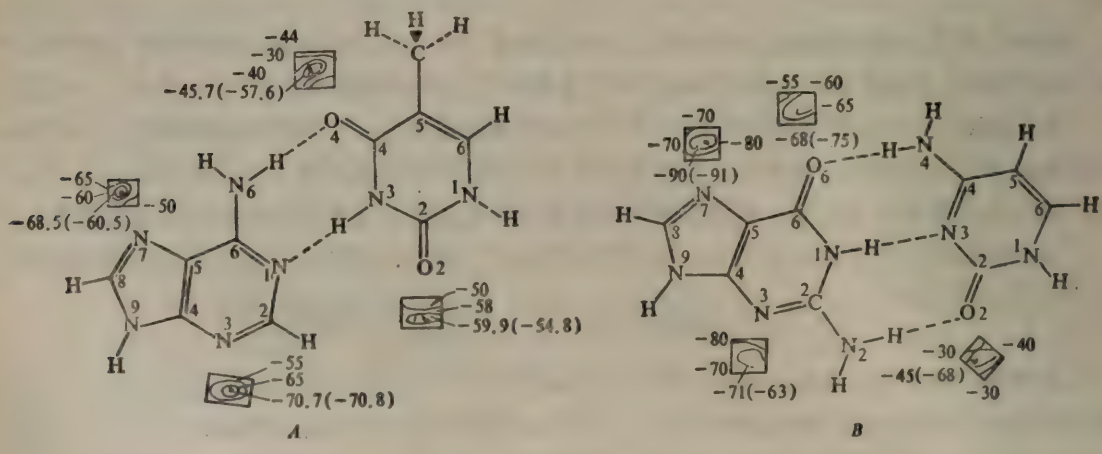


图 4-37 DNA 碱基平面上的分子静电势(*A*)A-T对，(*B*)G-C对

4.8.3 一段完整的B--DNA双螺旋的分子静电势

作为处理一段完整的B-DNA双螺旋模型的第一步，是先研究一小段螺旋。以G-C作为中心碱基对的情形为例(图4-38)，首先考虑螺旋的两条糖-磷酸根主链对中心G-C碱基对周围电势的影响(表4-38)，然后再与上、下碱基对的影响迭加(表4-39)，便可得到势能极小值的总数值。再次迭加这些小段螺旋基本单元的静电势，就可得到一段完整的双螺旋单位在空间任何一点的分子静电势。

表 4-38 G-C碱基对在不同情况时各个位置上的静电势极小值(kcal/mol)

位　置	单个碱基	配对碱基	在一段B-DNA中的配对碱基
N3(G)	−64.5	−73.7	−234.8
O6(G)	−76.6	−79.1	−204.4
N7(G)	−92.6	−92.9	−238.9
O2(C)	−70.6	−50.3	−217.8

表 4-39 中心G-C碱基对各位置受邻近碱基对影响后增减的静电势(kcal/mol)

位　置	↓G-C↑	↓$\overline{\text{G-C}}$↑	↓C-G↑	↓$\overline{\text{C-G}}$↑	↓T-A↑	↓$\overline{\text{T-A}}$↑	↓A-T↑	↓$\overline{\text{A-T}}$↑
N3(G)	−5.5	−3.2	7.0	−3.1	2.6	−7.4	−0.8	−6.9
O6(G)	−9.1	−12.3	−7.8	−0.7	−5.6	−2.5	−5.6	−5.4
N7(G)	−11.4	−7.9	−1.9	4.6	−4.6	0.8	−4.6	−1.5
O2(C)	−2.6	7.6	−2.7	−3.0	−6.9	2.1	−6.6	2.7

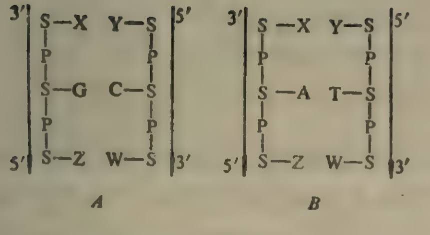


图 4-38 小段DNA螺旋模型

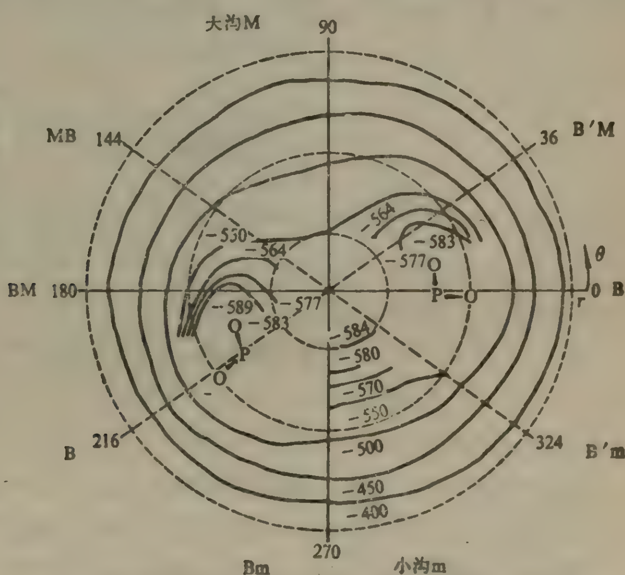


图 4-39 由糖-磷酸根主链产生的双螺旋在平面P上的等势线

根据这一思想，作者等[64,65]首先计算了仅由螺旋内外的糖-磷酸双螺旋主链所产生的分子静电势。计算时考虑了所有的磷酸根，但只考虑与各个计算平面相距最近的 4 个糖，计算范围是离螺旋轴 0.4 nm 至 1.6 nm。图 4-39 是平面 P 的计算结果，其中有几点是值得注意的：

(1) 由于磷酸根影响的总和效应，势的数值比单个磷酸根要大得多。

(2) 主链势在空间上扩展很远，递减较慢，数量上似乎与距离成反比，形状上具有明显的外圆内椭性。

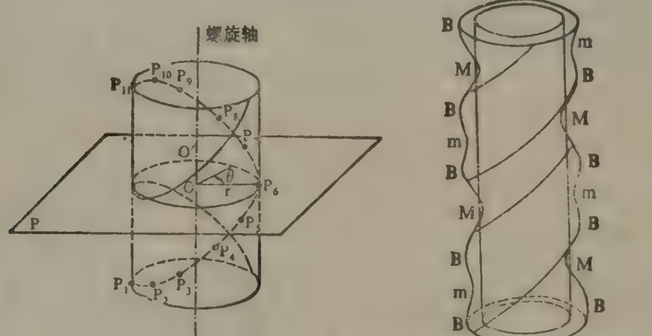
(3) 势能极小值的二个地方(－580 和－590 kcal/mol) 出现在大沟中部的两边。很清楚这是由于面向该沟的最邻近磷酸根上的氧阴离子所造成的。

离轴等距离时，由于指向小沟的糖氧原子的影响，使m方向电势优于M方向电势

(4) 相距 0.084 nm平行于平面P的一系列平面 P'、P'' 和P'''上的等势曲线与平面P非常相似。只是大沟中的势能极小值略有差别。分析上述各平面和各方向的电势，可以外推出由糖-磷酸根主链产生的三维等势面的一般特征。图 4-40(B)是 B-DNA 边界区域等势面的形状(包含了四小段完整的双螺旋)，它清楚地表现了分子的双螺旋特征。

在上述主链势计算的基础上，进而考虑引进G-C或A-T碱基对所引起的变化。在实际计算时，可以只考虑影响力最强的碱基对(平面P两边各二对)。图 4-41 是聚(dG・dC)的大沟和小沟以及在平面P上等势线的轮廓。从图中可以看出，大沟含N4(C)、N7(G)、O6(G) 而小沟含N2(G)、N3(G)和O2(C)。最引人注目的特征是聚(dG・dC)中的大沟有最强的势。与主链势相比，只是在距离螺旋轴较近的地方，等势线的形状和数值才因引入聚(dG・dC)而有所变化。

表 4-40 列出了碱基对上特别重要的几个活性位置的势极小值，并与自由碱基、核苷、核苷酸、单股螺旋和双股螺旋相应位置上的势值作比较。



A. 双螺旋模型　　B. B-DNA 外边界等势面的透视图

图 4-40

从表中可以看出，势极小值主要出现在碱基平面环氮和羰基氧附近，势能最小值通常在N7(G)。在碱基平面外，静电势极小值则出现在碱基的氨基、嘌呤的C8和嘧啶的C5═C6键等处。从表中可看出一个极其显著的特点是，随着体系的复杂化，静电势最小值的绝对值(深度)逐渐加大。当核苷酸经单螺旋时，增加尤为显著，这是由于磷酸根在碱基周围产生的强势及碱基本身电

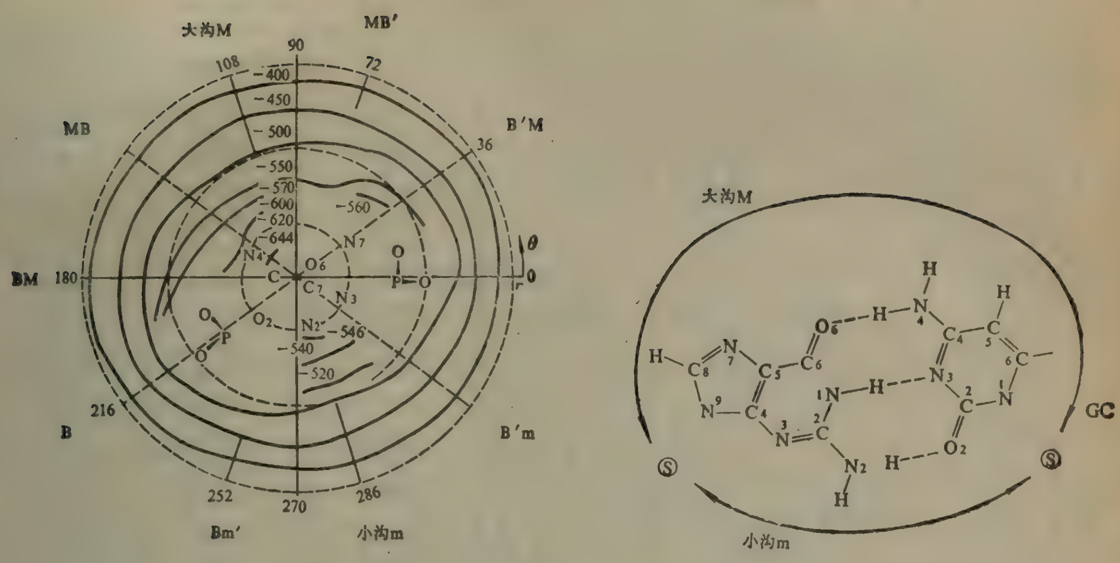


图 4-41

聚(dG·dC)中的主沟与小沟　聚(dG·dC)在 P 平面上的等势线

势迭加所致。从单螺旋过渡为双螺旋时，暴露于分子表面形成氢键的 N 1 (A)和 N 2 (C) 是上述电势深度变化规则的一个例外。同时，亲核中心的相对次序也往往会发生一些变化。

综上所述，有关核酸分子静电势的所有这些计算结果，在一定程度上，对于阐明生物分子的反应性等一些问题上，都有着重要的生物学意义。在本书后面的某些章节中还将予以讨论。不过，分子静电势法的一些不足之处，是应该予以注意的。

4.9　核酸的功能

核酸(尤其是脱氧核糖核酸)在生物系统中起着信息功能分子的重要作用。人们普遍认为，DNA 是绝大多数生物的原初遗传信息载体，不但通过复制、转录和转译过程可使其载有的遗传信息得到传递与表达，而且经过严格调控可使之有不同的表达，以致指令同一个受精卵严格有序地分化为不同类型的细胞，组成各种组织和器官，构成完整的生物个体。读者可从“量子遗传学”一章，了解到核酸作为信息高分子，在电子水平上对其功能所作的研究。

在这一节里，仅对以下四个问题作简要的介绍仍旧是必要的。当然，我们将极力避免与第七章在内容上重复。

4.9.1　作为遗传信息载体的 DNA[66-68]

DNA 作为遗传信息载体具有哪些特征呢？如何从电子水平上来看这些特征？下面我们就

表 4-40 核酸碱基上各亲核位置在不同情况下的静电势极小值(kcal/mol)

位 置	单个碱基	核 苷	核苷酸	单股螺旋	双股螺旋
N 7(G)	−88.2	−88.0	−146.8	−420.6	−682.6
N 3(G)	−55.6	−58.9	−100.9	−373.3	−670.0
N 3(A)	−62.6	−65.8	−107.1	−390.5	−668.8
O 2(T)	−56.2	−75.4	−110.5	−367.9	−662.8
O 6(G)	−75.0	−74.4	−115.2	−387.7	−654.4
N 7(A)	−67.1	−66.9	−127.2	−385.2	−650.2
O 2(C)	−82.3	−84.7	−122.6	−410.4	−645.1
C 8(G)5′	−8.0	−9.6	−73.3	−350.7	−630.3
C 8(G)3′	−8.0	−5.5	−68.2	−366.5	−623.4
N 2(G)3′	3.3	2.7	−29.5	−302.1	−623.1
N 2(G)5′	3.3	0.8	−31.6	−260.1	−622.6
O 4(T)	−55.0	−55.3	−97.2	−353.2	−611.7
C 8(A)5′	0.9	−5.1	−69.2	−341.9	−610.5
N 4(C)3′	−8.3	−8.2	−51.4	−299.6	−601.9
N 6(A)5′	−12.9	−14.0	−52.2	−313.3	−600.4
N 6(A)3′	−12.9	−13.0	−50.9	−302.7	−597.7
C 8(A)3′	0.9	1.3	−60.0	−337.3	−597.2
N 4(A)3′	−8.3	−9.3	−52.9	−322.2	−593.1
C5—C6(T)3′	1.0	−1.6	−69.3	−325.7	−591.6
C5—C6(C)5′	−3.2	−5.6	−70.8	−349.0	−584.1
C5—C6(T)5′	1.0	2.8	−61.1	−325.7	−583.6
C5—C6(C)3′	−3.2	−1.7	−64.2	−311.0	−569.0
N 3(C)	−85.7	−85.2	−126.7	−404.4	—
N 1(A)	−70.4	−71.3	−104.0	−367.4	—

来简略地讨论这些问题。

(1) 我们可以从 Wiener 所著的《控制论》一书中得到启发。Wiener 指出："许多存储信息的方法，共同具有一个重要的物理要素，这就是它们似乎都是高度量子简并系统。换句话说，都是振动方式很多但频率相同的系统。铁磁性物质的情况就是这样，具有很高介电常数的物质也是这样。因此这些物质特别适合用来作为储存信息的电容器。磷光现象也是高度量子简并性的"。Wiener 还指出："新陈代谢和生殖作用的许多问题都与量子简并性有关。"[69]

(2) 研究表明，DNA 是具有很高介电常数的物质，它的介电常数甚至超过目前所知最好的铁电材料——钛酸钡。在相变点附近钛酸钡的介电常数约为 18 000，而DNA 的介电常数则高达 120 000 以上[70]。是已知的介电常数最高的物质。

如果按照 Wiener 的上述观点，则 DNA 分子即是一种高度量子简并性物质。

(3) 在天然的 DNA 分子及其与蛋白质的络合物中，已证实有巨量的、强大相互作用的非偶电子云存在。

根据实验测定的结果，每一DNA 分子含有约 10 000～50 000 个非偶电子，约相当于平均每克 DNA 含有 10^{21} 个非偶电子。然而，正是由于 DNA 中大量非偶电子的存在，从而使之有可能

出现三线激发态(三重态)。

我们知道，三重态比单线态要稳定一些，寿命较长。一般，三重态的寿命比单线态增加约一百万倍，而单线态的能量相当于光谱的可见波段。这种能量的光量子为生物系统所利用似嫌太大。但是，当电子激发到三重态时，由于有一部分能量消耗在使电子自旋反向上，因而能量降低，磷光波段较长，接近红外线。这就为DNA大分子利用电子激发能来传递遗传信息提供了能量条件。

我们知道，生物分子中的高能磷酸键的键能约为10～12 kcal/mol(见4.2.5.2)，相当于红外区2～3 μm的光量子，氢键键能则相当于2.7 μm的光量子，也在这个范围内，而DNA双螺旋结构中正是含有大量的氢键。在2～3 μm的红外波段中，容易产生频率相同而方式又是多种多样的激发，这就是量子简并性。它使得DNA有可能进行多种多样、灵活而精确调谐的能量转移，并利用很少的能量来精确地、高效地储存和传递极其丰富的信息。

(4) DNA具有一定的产生共振转移的条件。共振转移是电子激发能的转移方式之一，又称感应转移或电磁偶联。产生共振转移的可能性(几率)取决于电子激发态的寿命。一个简并态被激发的寿命越长，激发能进行共振转移的几率也越大。

如前所述，DNA最重要和最可能出现的电子激发态是长寿命的三重态。根据量子力学的解释可设想：在DNA进行复制或转录时，很可能以进行共振转移的红外波段的三重态电子激发能作为"信道"向介质中的三磷酸核苷单体的原料发送信息，且DNA分子中特定的核苷酸也可能以特定频率的电子激发能发送信息，从而吸引与它配对的核苷酸单体，如A吸引T(或U)、G吸引C。

共振转移的特点是仅仅依靠两个分子之间的电磁相互作用而产生偶联。共振转移是一种长程作用，当分子间距离为50～10 nm时，它是传递电子激发能的主要方式。电子激发能可以通过电磁偶联在同一个大分子的各原子团之间进行转移，联想到DNA分子中的重复单位(脱氧核糖-磷酸)和碱基的共轭结构等等，很可能在DNA分子中所有基团都是如此，因而给予其中一个基团的能量，就可以被其它基团所吸收，整个高聚物分子形成一个统一的系统。DNA在复制或转录时，复制的起点或转录的起动子一旦被启动，整个DNA模板就被活化，复制或转录就一直进行到终止信号为止。

(5) DNA分子所能存储的信息量是极其巨大的。例如大肠杆菌(*E.coli*)一个基因组所含有的DNA，卷缩盘绕起来只有针尖的万分之一那样大，但其中所记载的遗传密码就大约相当于一部中篇小说。然而，在理论上，一个由1 000个碱基对组成的DNA分子片断，就可以有4^{1000}种不同的排列方式，这是一个极其巨大的天文数字，是电子计算机的存储材料所无法比拟的。

此外，有实验报道称DNA具有半导性(见4.5)，这也是它作为信息载体所具有的特征之一。

以上我们概括地讨论了DNA作为信息载体所具有的"重要物理要素"。从生物系统的角度看，作为遗传信息载体的DNA分子，除了自我复制以外，还具有以下一些特征：

(1) 细胞核中的DNA含量是恒定的。细胞中DNA含量分析结果表明，任何一种生物的

体细胞，其 DNA 含量都是固定的，不受外界环境、营养条件和细胞本身代谢状态的影响，这是符合作为遗传物质的要求的。此外，每个细胞中 DNA 含量与生物体的复杂性有近似的平行关系。可以想象，生物体越复杂，遗传信息量越大，这就需要有越多的遗传物质作为基础。

(2) 作为遗传信息载体必需在代谢上具有稳定性及其在细胞分裂中的完整传递。遗传学的研究表明，在遗传过程中，基因能保持独立性和稳定性。然而，一般人所认为的组成基因的 DNA 在代谢上是否是稳定的？应用同位素标记的 DNA，证明了细胞在分裂过程中 DNA 还经过分解和重新合成，就能完整地传于子代。因此，人们认为 DNA 很可能是构成基因的物质。

(3) 研究表明，当受到一些物理和化学因素作用时，可能增加有机体的突变率。一般说来，一些物理和化学因素能够引起 DNA 结构的改变，生成异常的 DNA 而引起突变率的增加。

4.9.2 核酸中电子激发能的转移方式[68]

以上我们从电子水平简要地概括了 DNA 作为遗传信息载体所具有的一些特征。须要注意的是，这些特征不是各自孤立地出现的，而是互为条件、互相影响和综合表现的统一的特性。这样，就使得核酸有可能进行灵活、而且振动组合方式多种多样的微观活性形式——电子激发能的转移，并且利用电子激发能来传递信息。

关于生物活细胞内电子激发能转移(包括激发电子本身直接转移)的可能方式，目前一般认为有：共振转移(电磁偶联)、重迭转移、络合物电荷转移、半导性能带的电子转移、激子转移、自由基、质子转移和氢键，等等。以下我们就简述遗传物质中电子激发能的转移方式与信息传递的关系(有关“生物系统中的能量转移”，我们还将较详尽地在第 11 章中讨论)。

(1) 共振转移是一种相同分子或不同分子之间的长距离的能量转移，是一种长程作用。电子激发能本身可以看成是一种振动，只要吸收能量的分子(受体)的振动电场是在激发分子(供体)的振动电场之中，而且供体的荧光光谱与受体的吸收光谱有重叠部分，激发分子的电磁振动就会引起吸收分子的同频共振，从而转移电子激发能。

由于核酸具有高度量子简并性，所以其中存在着大量的简并状态——内部共振。用结构化学的术语来说，核酸分子有着大量的共振结构式。它的电子离域性极强，分子的键型变异很大。由量子力学中变分法所取的波函数 ψ 值，对于核酸分子来说，是极其大量的。因此共振能很大，亦即能量降低很多，使分子趋向于很稳定的状态。这与作为遗传物质基础的核酸分子在世代传递中和代谢活动中所表现的稳定性是一致的。从电子水平来看，这是一个很有意义的事实。

核酸分子中各种原子团能象那样摆动、振动和转动。由于它的原子团是由磷酸酯键和氢键偶联起来的，它们能进行偶联振动子那样的动作，因此电子激发能就可以通过电磁偶联在同一个大分子的各原子团之间以及相邻的分子之间进行迁移(参见4.9.1)。

由于核酸分子中最重要的电子激发态是三线态，这对产生电子激发能的共振转移是十分有利的。因而核酸分子能量的共振转移很可能是由一个三线态转移到另一个三线态，以及它们之间的来回转移：

$$\begin{array}{ccc} A+B+\text{三线态 } C^{*}+\cdots & \rightleftharpoons & A+\text{三线态 } B^{*}+C\cdots \\ \uparrow & & \downarrow \\ & \text{三线态}A^{*}+B+C\cdots & \end{array}$$

（2）重叠转移是电子激发能的一种近程作用，当分子间的距离在 0.2—0.5 nm以下时（决定于电子轨道的大小），这种传能方式较易产生。

重叠转移现象也特别明显地表现在具有共轭双键系统的大分子内部的基团之间，前面已提及，DNA 具有广义共轭系统的性质，分子中的碱基都含有共轭双键系统，其中的 π 电子为整个系统所有，它们容易被激发并进行重叠转移形式的传能。

（3）在生物学过程中意义更为普遍的传能方式可能是络合物的电荷转移(参见第 3 章“电荷转移相互作用”一节)。

在 DNA 分子的双螺旋结构中，以特殊定向排列原子团是一个很重要的事实，其中的碱基基团可能发生电荷转移。在嘌呤和嘧啶的平行排列架中，实现电荷转移的条件比溶液中优越得多，而且电荷转移过程可沿这种分子架反复进行。1962 年张华就曾建议[71]，用一种形成电荷迁移络合物的正负电荷有序交错的双螺旋模型来解释 DNA 分子中大量非偶电子的产生，即

$$\begin{bmatrix} -\mathrm{Pr}^{+}-\mathrm{Pm}^{-}-\mathrm{Pm}^{+} \\ -\mathrm{Pm}^{-}-\mathrm{Pr}^{+}-\mathrm{Pr}^{-} \end{bmatrix} M/3 \qquad (4.17)$$

其中 Pr 表示嘌呤，Pm 表示嘧啶，M 是 DNA 分子碱基对的数目。

（4）由于电荷转移是单电子转移的过程。所以，络合物中若有一种物质具有晶体结构或者是分子集合体时，转移结果可以导致电子导电，亦即出现半导性。因此络合物电荷转移为半导性进入生物学开辟了道路。关于 DNA 的半导性问题，可参见 4.5。

（5）自由基是处于电子激发态的分子或分子的碎片。它们具有自由价，一般都有大的化学活性和特殊的电磁性质。DNA 分子中有着大量的相互作用的非偶电子云，处于三线态的电子自旋方向，就成为非偶电子。许多研究表明，处于三线态的分子往往是带有两个非偶电子的自由基。自由基在生命活动过程中具有相当重要的地位，是因为它的非偶电子与激发能的传递有密切关系。

实验表明，由自由基调节的反应能使某些致癌物质与 DNA 结合，且这种结合跟致癌活性有关。经化学诱变的核酸比通过加热或变性剂作用的核酸自由基增加较多。变性 DNA 甚至更具有反应活力。迄今为止，已经有不少人研究了 DNA 的改变和自由基的形成、转变间的关系[72]。

一种可能的情形是：在生活状态下活细胞及其中的 DNA 大分子有可能利用某种稳定的、长寿命的复杂自由基来转移能量和传递信息[68]。

（6）核酸具有很高的介电常数，这对能量和信息传递也有重要意义。在具有高介电常数的晶体中，激发电子的轨道可以笼罩许多分子，因此能量沿晶格的迁移表现为一种“正穴”(空穴)传递，而与正穴相应的激发电子则离正穴尚有一定距离，它们共同组成一个中性激发单位，叫做激子(exciton)(见第 11 章)。

Reid 曾指出，在生物体内普遍存在大分子的有序但非均质的系统，激子转移可能有相当的重要性。DNA 分子本身具有高度有序的结构，而且 DNA 的介电常数高达 120 000 以上，这就为激子转移创造了条件，并且可以使激子保存较久。

（7）研究表明，生物学的许多基本过程不仅与电子的转移相联系，而且也与质子的转移有关。Kirkwood 指出，生物高分子如核酸和蛋白质等，分子表面上有大量的质子。这些质子不

断地运动着，质子转移的同时也就伴随着激发能的传递。DNA 大分子与周围分子的相互作用，也可能受到质子运动的影响。质子运动可以与邻近类似的分子相呼应，构成能量转化和传递的系统，使 DNA 能够利用这种传能方式来传递信息。

我们知道，以质子相互联系而形成的化学键即氢键。DNA 分子的两条单链以氢键维系着而形成严格有序的双螺旋结构。氢键中质子的摆动也能转移能量。在 DNA 复制和转录过程中，作为反应原料的三磷酸核苷中高能磷酸键的能量就很容易转移到氢键上而与模板链上的核苷酸进行配对，形成新的 DNA 或 mRNA 分子。此外，氢键两端的 π 电子云有足够的重叠，π 电子可以横越氢桥而实现非定域化，有助于实现激发能的共振转移和半导性能带的电子转移。

Lowdin 曾提出 DNA 分子中构成氢键的质子具有隧道效应。形成氢键时，质子通常在氢供体的旁边，是一种较稳定的状态，而转移到对称的氢受体旁边时，也是一种较稳定的状态。用量子力学的术语来说，就是氢键具有对称的双势阱(参见 3.2.2)。

有关质子隧道效应引起基因突变的问题，在3.2.2中已经作过较详尽的讨论，这里不再叙述。

4.9.3 DNA 的复制

DNA 必须通过自我复制才能把遗传信息传递给子细胞和后代个体。前面，我们讨论了 DNA 作为遗传信息载体所具有的一些特征。那么，DNA 具有信息结构的基础究竟是什么？

根据迄今为止的研究结果看，显然多聚磷酸和糖的结合体不带有信息，那么信息可能就是来源于碱基部分。大量的事实确实支持了这样的推测，即核酸的自我复制就是以专一的碱基对的互补性为基础的。如果说 DNA 的复制是基于专一的碱基的互补性，那么，DNA 复制中碱基对的这种互补性又是由什么因素控制的呢？人们曾提出过氢键控制说，酶控制说等。Rein 等[7a]对形成互补链的各种物理因素进行了量子力学计算。结果显示，构象能在碱基对的互补性上起着决定作用。

计算时采用的模型是：先有一股母链，然后产生出一条互补的子链。由于计算量很大，因而在实际计算时，只计算了在母链上有两个相邻的核苷酸（各带有特定的碱基 1 和 2，见图 4-42）子链上已有了一个与母链上碱基 1 互补的碱基 3 的情况下，当再加上第四个碱基时，形成 DNA 双螺旋结构的构象适应能。为说明问题起见，Rein 等作了②-④ 碱基对是正常的(即互补的)和

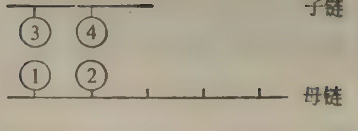


图 4-42

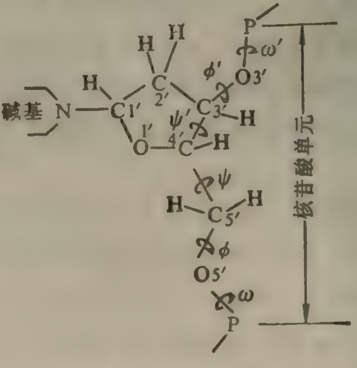


图 4-43　多核苷酸的主要扭转角

非互补的这两类不同情况的构象能计算，结果见表 4-41，表中有关的各个扭转角，可见图 4-43。

由表 4-41 可以看出，互补构型比非互补构型（如 G-U）的构象能要低得多。

表 4-41 构象能（kcal/mol）（高于最低值）

	正　常	U-U	G-U
χ_{C-N}/ψ	4.5	4.5	6.0
χ_{C-N}/ϕ'	4.5	3.6	13.0
ω/ω'	3.0	3.0	3.0
ω/ϕ	0.5	1.0	1.0
ϕ/ψ	0.5	2.0	1.0
总	13.0	14.1	24.0

用计算构象能的上述模型，进一步考虑在"复制络合体"形成过程中，在已有一对互补碱基的基础上，再形成一对碱基所引起的焓变 ΔH 与自由能改变 ΔG：

$$\Delta H = E_{氢键} + E_{溶剂效应} + E_{堆积} + E_{构象} \tag{4.18}$$

$$\Delta G = \Delta H - T\Delta S \quad (\Delta S_{实}是实验估计的\ \Delta S)$$

由表 4-42 可以看到，非互补的错误排列构型的ΔG 很大，互补性的决定因素是形成"复制络合体"的ΔG。组成不同的 DNA 的 ΔS，一般而言，具有大致相同的值，故而影响 ΔG 的是 ΔH，而在 ΔH 所包含的诸因素中，构象能又是决定因素，这就是 Rein 等人的结论。

表 4-42 "复制络合体"的热力学量（kcal）

构　　型	$\Delta H_{计算}$	$\Delta H_{实}$	$\Delta G_{计算}$	$\Delta G_{实}$
↑G-C↓ G-C	−10.8	−13.7	−2.0	−4.8
↑A-T↓ G-C	−3.6	−5.9	+0.1	−2.1
↑G-G↓ A-T	−5.1	−5.9	−1.4	−2.1
↑A-T↓ T-A	−3.5	−6.5	+1.3	−1.6
↑U-U↓ G-C	+1.1		+11.0	
↑A-T↓ U-U	+1.4		+11.4	
↑G-U↓ G-C	−2.3		+7.6	
↑U-U↓ T-A	+8.3		+10.2	

按照氢键控制说，在 DNA 的碱基配对中起着决定性作用的应是氢键。然而，实际上当碱基发生异构化作用时，也同样可以配对，但这时已不再符合原来 A-T、G-C 的配对规律。

如果根据异构化作用所需要的不很大的能量计算，则产生错误配对从而产生突变的几率为 10^{-4}—10^{-5}。但实际上自然突变的几率仅为 10^{-8}—10^{-11}，远远小于计算结果。由此可见，在碱基

配对中不能把氢键看作是唯一的因素，甚至并不是主要因素[43]。

4.9.4 DNA 的甲基化作用与基因功能

DNA 甲基化问题的研究可能对深入了解基因表达的分子基础是有价值的。

关于 DNA 甲基化作用的生物学意义，若干年来，不少人认为 DNA 的甲基化与基因调控有关，并十分引人注目，但实验事实却捉摸不定[74]。大量的观察和研究提示：在脊椎动物中，DNA 的甲基化以这样或那样的方式贯穿于重要的生命过程，而且甲基化模式具有遗传连续性、种属和组织特异性[75-76]。总之，一般认为，在真核生物中 DNA 的甲基化具有多种生物学功能，如对转录的控制，染色体结构的维持，DNA 的修复，癌变，以及在某些系统中还可防止 DNA 的酶促降解[77-78]。

我们知道，正常碱基的配对是 A 和T、G 和C。但是，当 DNA 双螺旋结构中两条互补的多聚脱氧核苷酸链上的胞嘧啶碱基在甲基化酶的催化下，经甲基化反应转换成 5-甲基胞嘧啶后，在 DNA 大沟的暴露部位引入了甲基基团，由于甲基基团的立体效应和电子效应的影响，使乳糖阻遏物，组蛋白和激素受体等等蛋白与 DNA 的结合受到大沟变化的影响[74]。研究事实已经说明，5-甲基胞嘧啶能够深刻地影响蛋白质与DNA 的结合[74]。一些人认为，在一定条件下，5-甲基胞嘧啶能够发生脱氨基反应生成胸腺嘧啶（图 4-44），从而有可能导致 G-C→A-T 的转变引起基因突变[79,80]。

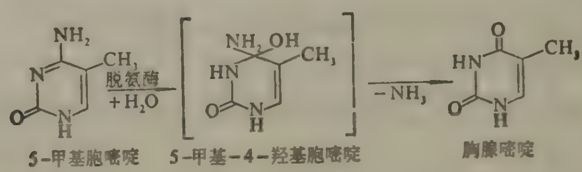


图 4-44

Coulondre 等人[81,82]根据在原核生物中的某些实验结果指出，DNA中胞嘧啶C 5 位的甲基化与高的基因突变率直接相关，突变的“热点”就在 5-甲基胞嘧啶。可是，根据作者用分子轨道法所作的理论计算，却得到了一些可争议的结果[50,82]。首先，计算结果显示：胞嘧啶甲基变成5-甲基胞嘧啶后，除C 5 位的电荷密度、自由价等数值有所减少，C 6 和C 4 位则有所增加外，其余相应位置的净电荷、自由价等指数并未发生明显的改变（表 4-43），且 G-C 碱基对与G-m^5C 碱基对氢键的键级完全相同，前者为 0.035，0.037，0.035，后者亦是 0.035，0.037，0.037。

在 m^5C 水解脱氨基反应的中间物——5-甲基-4-羟基胞嘧啶中，C 4—N 7 键的键级为 0.502，明显地大于C 4—O 9 键的键级 0.344。

m^5C 的 $\sum S_r^E - \sum S_r^N$ 大于C（表 4-44）。在碱基配对（G-m^5C）的情形下尤为明显。

计算结果还提示了以下的可能性：

(1) C 和 m^5C 两者有着十分相近的反应性，即胞嘧啶C 5 位引入甲基基团后对其分子的行为不致产生明显的影响。

表 4-43

	HMO							EHMO	CNDO/2
	键	级	自由价		S_r^E	S_r^N	净电荷	净电荷	净电荷
胞嘧啶	1—2	0.375	1	0.511	1.422	−0.357	0.361	−0.898	−0.174
HOMO: 0.595	2—3	0.433	2	0.146	0.307	−0.643	0.204	1.784	0.425
LUMO: −0.795	3—4	0.636	3	0.346	1.709	−0.474	−0.438	−1.466	−0.344
	4—5	0.525	4	0.101	0.467	−1.009	0.172	1.228	0.329
	5—6	0.758	5	0.449	1.435	−0.572	−0.169	−0.246	−0.165
	6—1	0.529	6	0.445	0.700	−1.242	0.165	0.626	0.189
	2—8	0.778	7	0.944	1.758	−0.197	0.197	−1.089	−0.229
	4—7	0.470	8	0.222	1.016	−0.283	−0.492	−1.332	−0.411
5-甲基胞嘧啶	1—2	0.377	1	0.517	1.479	−0.354	0.358	−0.901	−0.172
HOMO: 0.530	2—3	0.433	2	0.145	0.315	−0.642	0.203	1.774	0.423
LUMO: −0.796	3—4	0.637	3	0.344	1.722	−0.475	−0.436	−1.465	−0.310
	4—5	0.515	4	0.111	0.489	−1.011	0.166	1.205	0.321
	5—6	0.759	5	0.257	1.479	−0.567	−0.123	−0.076	−0.125
			6	0.464	0.813	−1.217	0.138	0.574	0.171
	6—1	0.520	7	0.945	1.776	−0.199	0.198	−1.103	−0.232
	2—8	0.776	8	0.224	1.043	−0.283	−0.493	−1.333	−0.410
	4—7	0.470	9	1.520	0.542	−0.441	−0.047	−0.267	−0.003
	5—9	0.212	10	0.932	0.632	−0.520	0.953		
5-甲基-4羟基胞嘧啶	1—2	0.370	1	0.527	1.499	−0.333	0.348		
HOMO: 0.567	2—3	0.372	2	0.181	0.278	−0.674	0.231		
LUMO: −0.774	3—4	0.483	3	0.559	1.374	−0.269	0.313		
	4—5	0.503	4	−0.100	0.398	−1.082	0.252		
	5—6	0.751	5	0.266	1.500	−0.555	−0.143		
	6—1	0.517	6	0.464	0.723	−1.294	0.167		
	2—8	0.810	7	0.912	1.660	−0.214	0.211		
	4—7	0.502		0.190	0.866	−0.302	−0.463		
	4—9	0.344		0.656	0.955	−0.094	0.097		
	5—10	0.212	1	1.520	0.542	−0.441	−0.047		
			1	1.732	0.634	−0.519	0.034		
G-C对(三个氢键)	1—2	0.383	1	0.504	1.418	−0.358	0.363	−0.896	−0.142
HOMO: 0.288	2—3	0.437	2	0.150	0.296	−0.678	0.232	1.793	0.426
LUMO: −0.782	3—4	0.588	3	0.351	1.596	−0.413	−0.509	−1.479	−0.349
	4—5	0.524	4	0.094	0.442	−1.033	0.184	1.252	0.317
	5—6	0.759	5	0.449	1.433	−0.574	−0.169	−0.241	−0.174
	6—1	0.527	6	0.446	0.677	−1.267	0.173	0.592	0.195
	2—8	0.762	7	0.852	2.063	−0.254	0.249	−1.065	−0.229
	4—7	0.526	8	0.204	0.925	−0.270	−0.524	−1.358	−0.459
	9—10	0.410	9	0.472	1.668	−0.264	0.358	−0.951	−0.188
	10—11	0.386	10	0.171	0.376	−0.704	0.219	1.285	0.371
	11—12	0.624	11	0.219	1.793	−0.492	−0.180	0.288	−0.110
	12—13	0.486	12	0.137	0.999	−0.731	0.025	1.063	0.232
	13—14	0.627	13	0.302	2.345	−0.376	−0.432	−1.453	−0.333
	14—9	0.495	14	0.102	0.662	−0.936	0.169	1.654	0.386
	10—19	0.764	15	0.164	1.322	−0.550	−0.289	−1.503	−0.252
	14—18	0.508	16	0.463	1.519	−0.852	0.010	1.181	0.156
	11—15	0.502	17	0.407	1.289	−0.335	0.404	−0.901	0.082
	15—16	0.478	18	0.871	2.389	−0.189	0.227	−1.083	−0.259
	16—17	0.522	19	0.201	1.090	−0.309	−0.509	−1.331	−0.420
G-m^{5}C对	1—2	0.386	1	0.510	1.474	−0.355	0.360	−0.898	−0.105
(三个氢键)	2—3	0.437	2	0.149	0.304	−0.676	0.230	1.781	0.442
HOMO: 0.287	3—4	0.589	3	0.350	1.602	−0.414	−0.507	−1.477	−0.359
LUMO: −0.78	4—5	0.514	4	0.104	0.465	−1.034	0.177	1.230	0.308
	5—6	0.749	5	0.257	1.477	−0.570	−0.123	−0.068	−0.131
	6—1	0.518	6	0.465	0.788	−1.241	0.146	0.512	0.179
	2—8	0.760	7	0.853	2.091	−0.256	0.250	−1.073	−0.231
	4—7	0.526	8	0.205	0.916	−0.270	−0.526	−1.360	−0.488
	9—10	0.410	9	0.171	1.668	−0.264	0.358	−0.951	−0.148
	10—11	0.386	10	0.171	0.376	−0.704	0.219	1.285	0.393
	11—12	0.624	11	0.219	1.791	−0.492	−0.180	0.288	−0.109
	12—13	0.486	12	0.137	0.999	−0.731	0.025	1.063	0.215
	13—14	0.627	13	0.301	2.316	−0.376	−0.432	−1.453	−0.349
	14—9	0.495	14	0.102	0.662	−0.936	0.169	1.654	0.391
	10—19	0.764	15	0.164	1.322	−0.550	−0.289	−1.503	−0.305
	14—18	0.508	16	0.463	1.519	−0.852	0.010	1.181	0.174
	11—15	0.502	17	0.407	1.289	−0.335	0.404	−0.901	−0.019
	15—16	0.748	18	0.871	2.390	−0.189	0.227	−1.083	−0.265
	16—17	0.522	19	0.201	1.090	−0.309	−0.509	−1.331	−0.463
	5—20	0.212	20	1.520	0.542	−0.441	−0.047	−0.265	−0.025
			21	1.732	0.631	−0.521	−0.037	0.423	0.111

表 4-44

化合物	$\Sigma S_r^E - \Sigma S_r^N$
G·m⁵C碱基对	14.261
G·C碱基对	13.715
G	8.754
5-甲基-4-羟基胞嘧啶	4.653
A	4.600
m⁵C	4.582
T	4.324
C	4.307

(2) m^5C水解脱氨基成T，从而引起G-C向A-T转化的反应是不大可能进行的(表4-43和表4-45)，水解脱氨基的反应速度随胞嘧啶C5位的甲基化而降低。有实验证明，脱氧胞苷水解脱氨基的反应速度大于5-甲基脱氧胞苷；胞嘧啶和5-甲基胞嘧啶两者都是胞嘧啶脱氨基酶的底物，且具有相同的K_m值[84-86]。

表 4-45 氨基的电子特征

化合物	氨基N原子		跟氨基相联的C原子	
	电荷	自由价	电荷	自由价
G	1.803	0.933	0.807	0.124
C	1.803	0.944	0.828	0.101
m⁵C	1.802	0.945	0.834	0.111
5-甲基-4-羟基胞嘧啶	1.789	0.912	0.748	0.100
G(在G—C对中)	1.773	0.906	0.831	0.102
G(在G—m⁵C对中)	1.773	0.906	0.831	0.102
C(在G—C对中)	1.751	0.888	0.816	0.094
C(在G—m⁵C对中)	1.750	0.888	0.823	0.104

不仅如此，作者还计算了胞嘧啶C5位甲基化作用的势能变化以及G-m^5C碱基对第一个氢键体系中质子移动的势能。结果显示，C和m^5C之间没有明显差别；G-m^5C对第一个氢键体系与G-C对一样，不存在双势阱。

为了探讨甲基化作用对DNA双螺旋结构稳定性的影响，作者计算了m^5C对碱基堆积相互作用的影响。按照分子间相互作用理论，在计算时，无论堆积相互作用还是共平面相互作用，静电项和极化项均按同样公式计算，而对色散-排斥项则分别采用Lennard-Jones的"6-12"势公式和经改进的Kitagorodsky公式(参见3.2.1)。计算的碱基堆积相互作用能见表3-4，而邻接碱基对的相互作用能列于表4-46。

计算结果表明，在单股核苷酸链的情形下，碱基堆积相互作用能中色散能的贡献是重要的。胞嘧啶经甲基化作用成为5-甲基胞嘧啶后，一般说来，对DNA双螺旋结构起着稳定的作用，如

表 4-46 邻接碱基对的相互作用能(kcal/mol)

	静电能	极化能	色散能	排斥能	总能量
$\uparrow\begin{matrix}C\cdot G\\ G\cdot C\end{matrix}\downarrow$	−8.704	−2.572	−8.407	5.186	−14.497
$\uparrow\begin{matrix}m^5C\cdot G\\ G\cdot C\end{matrix}\downarrow$	−8.776	−2.590	−8.442	5.186	−14.622
$\uparrow\begin{matrix}C\cdot G\\ G\cdot m^5C\end{matrix}\downarrow$	−8.822	−2.619	−10.248	6.253	−15.436
$\uparrow\begin{matrix}m^5C\cdot G\\ G\cdot m^5C\end{matrix}\downarrow$	−8.894	−2.636	−10.281	6.251	−15.560
$\uparrow\begin{matrix}G\cdot C\\ G\cdot C\end{matrix}\downarrow$	−7.356	−2.559	−8.091	4.983	−13.023
$\uparrow\begin{matrix}G\cdot m^5C\\ G\cdot C\end{matrix}\downarrow$	−7.394	−2.578	−8.144	4.983	−13.133
$\uparrow\begin{matrix}G\cdot C\\ G\cdot m^5C\end{matrix}\downarrow$	−7.406	−2.607	−9.833	6.015	−13.881

$\uparrow\begin{matrix}G\\ C\end{matrix}\rightarrow\uparrow\begin{matrix}G\\ m^5C\end{matrix}$,其总能量由−8.822 降低至−10.546 kcal/mol。在双股核苷酸链的情形下,m^5C 取代C后,碱基对的堆积总能量明显降低,如由$\uparrow\begin{matrix}C\cdot G\\ G\cdot C\end{matrix}\downarrow\rightarrow\uparrow\begin{matrix}C\cdot G\\ G\cdot m^5C\end{matrix}\downarrow$,其总能量由−14.497 降低至−15.436 kcal/mol。

由上述计算结果可以看到,DNA 的甲基化作用对 DNA 双螺旋结构起着一定的稳定作用,这与其它一些人的研究结果相一致[76~78]。

业已知道,原核生物的基因具有连续的编码序列,而高等真核生物基因则是以复杂的内含子-外显子结构为特征的。在动物细胞中,甲基化的胞嘧啶常常相对大量地存在于不转录的随体 DNA 顺序中。这些随体 DNA 的重复数目在不同的细胞类型中的波动,常常掩盖了 DNA 甲基化中较有意义的变化。

概括地,由于DNA 双螺旋结构中某些区域的胞嘧啶碱基被甲基化为 5-甲基胞嘧啶后,增加了这些区域双螺旋结构的稳定性。从这个意义上讲,DNA 的甲基化也许是导致不转录的间隔序列产生的原因之一。

如前所述,尽管有实验表明,在一些细菌和噬菌体中碱基取代的热点出现在 m^5C,进而认为由于碱基甲基化成稀有碱基(如 m^5C)后,细菌限制性内切酶不能识别和排除 DNA 中甲基化了的这些碱基,从而可能引起 G-C→A-T 的转化,导致基因突变率的上升。但是,在高等生物的DNA中,m^5C作为一种稀有碱基的存在却是它们的一个特点,至于G-C→A-T的转化(由m^5C导致的转化)能否发生,则是一个有待作进一步研究的问题(从我们初步的理论计算结果看,由 m^5C 引起的 G-C→A-T 转化的可能性是不太大的)。

在一些实验事实仍旧是捉摸不定的情形下,从量子生物学的角度就"DNA 的甲基化作用与基因功能"作些理论计算和研究,在作者看来是必要的和有益的。我们获得的初步结果也支持了这一点。

参考文献

[1] 乐树云，江寿平：科学通报，29:302 (1984).
[2] 乐树云，江寿平：生物物理学报，2:76 (1986).
[3] 刘次全等：动物学研究，5:79 (1984).
[4] 宣建诚，江寿平：中国科学，(7):681 (1986).
[5] 江寿平，宣建诚：中国科学，(6):609 (1987).
[6] 江寿平，宣建诚：科学通报，32:1182 (1987).
[7] Pullman, B. et al.: "Quantum Biochemistry", (Wiley Interscience, New York. London) (1963).
[8] Pullman, B. et al.: "Progress in Nucleic Acid Research and Molecular Biology", Davidson, J. N.(Ed)(Academic Press, New York, London),(1969).
[9] Praga, S. et al.: "Biomolecular Conformation Theory", (Elsevier Scientific Publishing Company, New York), (1978).
[10] 吉林大学生物系生物化学教研室："量子生物学"讲义。
[11] 福井谦一著，李荣森译：《化学反应与电子轨道》，科学出版社，(1985).
[12] Mely, B. et al.: Theoret. Chim. Acta., 13:278 (1969).
[13] Weinstein, I. B. et al.: Science, 193:592 (1976).
[14] Savin, F. A. et al.: Inter. J. Quantum Chem., 16:825 (1979).
[15] Borodavkin, A. V. et al.: Itoginauki Tekh. Mol. Biol., 14:3 (1977).
[16] Snyder, L. G. et al.: J. Chem, Phys., 53:256 (1970).
[17] 永田亲义著，陶宗晋、江寿平译：《量子生物学入门》，上海科学技术出版社，(1979).
[18] Mezey, A. R. et al.: Inter. J. Quantum Chem. QBS., (9):305(1982).
[19] Pullman, B. et al.: Proc. Natl. Acad. Sci. U. S. A., 55:904 (1966).
[20] Kwiatkowski, J. S. et al.: Inter. J. Quantum Chem., 15:499 (1979).
[21] Boyd, D. B. et al.: J. Theor. Biol., 25:403 (1969).
[22] Alving, R. E. et al.: ibid., 34:199 (1972).
[23] 江寿平、宣建诚：科学通报，30:1819 (1985).
[24] Pullman, B. et al.: Radiation Res., Suppl. 2:160 (1960).
[25] Alving, R. E. et al.: J. Theor. Biol., 34:199 (1972).
[26] A. 圣-乔奇著，林克椿译：《生物能力学》，科学出版社，(1962).
[27] Epp, A. et al.: J. Am. Chem. Soc., 80:724 (1958).
[28] Sachesenbeimer, W. et al.: J. Mol. Biol., 114:37 (1976).
[29] Ladik. J. et al.: Inter. J. Quantum Chem. QBS., 10:73 (1983).
[30] Otto, P. et al.: J. Chem. Phys., 78:4547(1983).
[31] Clementi, E. et al.: Inter. J. Quantum Chem. QBS., 9:213 (1982).
[32] 乐树云、江寿平：生物化学与生物物理学报 17:557 (1985).
[33] 乐树云、江寿平：生物化学与生物物理学报，17:664(1985).
[34] 乐树云、江寿平：生物物理学报，4:227 (1987).
[35] Beukers, R. et al.: Biochem. Biophys. Acta. 41:550 (1960).
[36] Smith, K. C. et al.: "Molecular Photobiology", (Academic Press, New York), (1969).
[37] Setlow, R. B. et al.: J. Mol. Biol., 17:237 (1966).
[38] Rahn, R. O. et al.: Biophys. J., 9:510 (1969).
[39] Setlow, R. B. et al.: Proc. Natl. Acad. Sci. U. S. A., 48:1250 (1962).
[40] Pullman, B. et al.: Photochem. Photobiol., 7:525 (1968).
[41] Fukui, K. et al.: Bull. Chem. Soc. Japan, 34:1178 (1961).
[42] Nagata, C. et al.: J. Theor. Biol., 9:357(1965).
[43] 程极济等：《生物物理学》，人民教育出版社，(1982).
[44] Ladik, J.: "Electronic Structure of Polymers and Molecular Crystals", (Plenum Press, New York, London), (1975).
[45] Ladik, J. et al.: Inter. J. Quantum Chem. QBS., 2:133 (1975).

[46] Hosur, R. V. et al.: Inter. J. Quantum Chem., 20:33 (1981).
[47] Otto, P. et al.: J. Chem. Phys., 80:5249 (1984).
[48] Ladik, J.: Inter. J. Quantum Chem., 9:563 (1975).
[49] 乐树云、江寿平：分子科学学报，1:87 (1982).
[50] Liu Ciquan, et al.: J. Mol. Sci., 4:301 (1986).
[51] Pullman, B. et. al.: Theor. Chim. Acta, 50:317 (1979).
[52] 吴伟雄和江寿平等：生物化学与生物物理学报，13:495 (1981).
[53] Soukhodub, L. F. et al.: Nature, 264:425 (1976).
[54] Ronnaccors, R. et al.: Theor. Chim. Acta, 36:339 (1975).
[55] Ronnaccors, R. et al.: Theor. Chim. Acta, 24:51 (1972).
[56] Perahia, D. et al.: Theor. Chim. Acta, 48:263(1978).
[57] Pullman, A. et al.: Chem. Phys. Lett., 32:233 (1975).
[58] Pullman, A. et al.: Chem. Phys. Lett., 41:205(1976).
[59] Berthod, H. et al.: Theor. Chim. Acta, 47:59 (1978).
[60] Lavery, R. et al.: Nucleic Acids Res., 8:1061 (1978).
[61] Zakrzewska, C. et al.: Inter. J. Quantum Chem., 16:395 (1979).
[62] Perahia, D. et al.: Inter. J. Quantum Chem. QBS., 7:323 (1980).
[63] Goldblum, A. et al.: Inter. J. Quantum Chem., 15:121 (1979).
[64] 邵俊等：大自然探索，(3):19 (1983).
[65] 邵俊等：大自然探索，(1):85 (1984).
[66] 王身立：自然杂志，(3):886 (1980).
[67] 刘次全等：北京轻工业学院学报，2:43(1985).
[68] 王身立,«控制论与遗传进化问题»，湖南科学技术出版社，(1985).
[69] N. 维纳著，郝季仁译：«控制论»（第二版），科学出版社，(1963).
[70] 陈春先：科学通报 (10):30 (1961).
[71] 张华：北京大学学报（自然科学版），(1):31 (1961).
[72] 刘次全：国外医学分子生物学分册，2:76 (1980).
[73] Rrin, R. et al.: Inter. J. Quantum Chem. QBS., 2:145 (1975).
[74] Razin, A. et al.: Science, 210:4470 (1980).
[75] Desrosiers, R. C. et al.: Proc. Natl. Acad. Sci. U. S. A., 76:3839 (1979).
[76] Ehrlich, M. et al.: Nucleic Acids Res., 10:2709 (1982).
[77] Razin, A. et al.: "DNA Methylation, Biochemistry and Biological Significance", (Springer-Verlag, New York, Berlin, Heidelberg, Tokyo), (1984).
[78] Cooper, D. N. et al.: Human Genetics, 64:315 (1983).
[79] Шабарова, З. А. И ДР.: "Химия Нуклеиновых Кислот и их Компонентов," (Издательство "Химия", Москва), (1978).
[80] Scarano, E.: Adv. Cytopharmacol., 1:13 (1971).
[81] Coulondre, C. et al.: Nature, 274:775 (1978).
[82] Hblliday, R. et al.: Science, 187:226 (1975).
[83] 刘次全等：分子科学学报，4:501 (1984).
[84] Liu ciquan, et al.: J. Mol. Sci., 4:301 (1986).
[85] Hoffee, P. A. et al.: "Methods in Enzymology", Vol. LI, (Acad. Press, New York, San Francisco, London), (1978).
[86] Ipata, P. L. et al.: Biochemistry, 10:2470 (1971).

刘次全
王身立
谭载友
王　莹

第5章　蛋白质的电子结构

蛋白质是一类重要的生物高分子。不同的蛋白质具有不同的生理功能。

蛋白质是各种 α-氨基酸通过肽键，即酰胺键联接而成的长链分子，这种长链称为肽链。链中相当于氨基酸的单元结构称为残基。肽链中氨基端称为氨端或N端，羧基端称为羧端或 C 端。有的蛋白质还包含有非肽链结构的其它成分，这种成分称为配基或辅基。图 5-1 标出了上述诸定义。

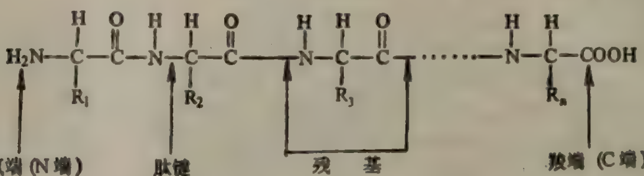


图 5-1　蛋白质分子肽链图

近年来，蛋白质研究已经发展到了结构-功能关系及蛋白质-蛋白质(或其它生物高分子、药物分子等)相互关系阶段。血红蛋白的“分子肺”功能的阐述、血红蛋白与镰刀状贫血分子病关系的揭示等等是从分子水平研究蛋白质的成功范例。从电子水平研究蛋白质结构-功能关系则是量子生物学的艰巨任务。

我们认为，对于分子结构，尤其是生物分子的结构信息，不仅应当包括在分子层次上由实验直接获得的基本结构信息(如蛋白质分子的各级结构所表达的)，也应当包括对基本结构信息进行理论加工所产生的各类理论结构信息，特别是反映分子中电子运动的微观结构信息和能够形象地再现大分子高级结构并有其明确的物理、化学、生物学意义的精细结构信息。理论生物学、理论化学的发展和计算机的运用为获取理论结构信息创造了良好的条件。量子生物学工作者正是沿着这条路线对氨基酸、多肽进而对蛋白质等大分子开展研究的。

在这章里，我们将对研究氨基酸、多肽和蛋白质的电子结构和溶液结构时所采用的基本思想、基本方法以及基本结果作一些介绍。

5.1　氨基酸的电子结构和氨基酸溶液结构的模拟

含有氨基和羧基的有机化合物都叫做氨基酸。在蛋白质中出现的氨基酸是一种特殊的氨基酸，其氨基和羧基都在 α-碳原子上，故称 α-氨基酸，如图 5-2 左上方所示。其中 R 称侧链，R 不同，就是不同的氨基酸了。

作为蛋白质组成成分的氨基酸共有 20 种。它们的名称和结构列于表 5-1。为了表示蛋白质和肽的结构，每个氨基酸都有自己的符号(取它英文名字的前三个字母)。一个蛋白质分子往

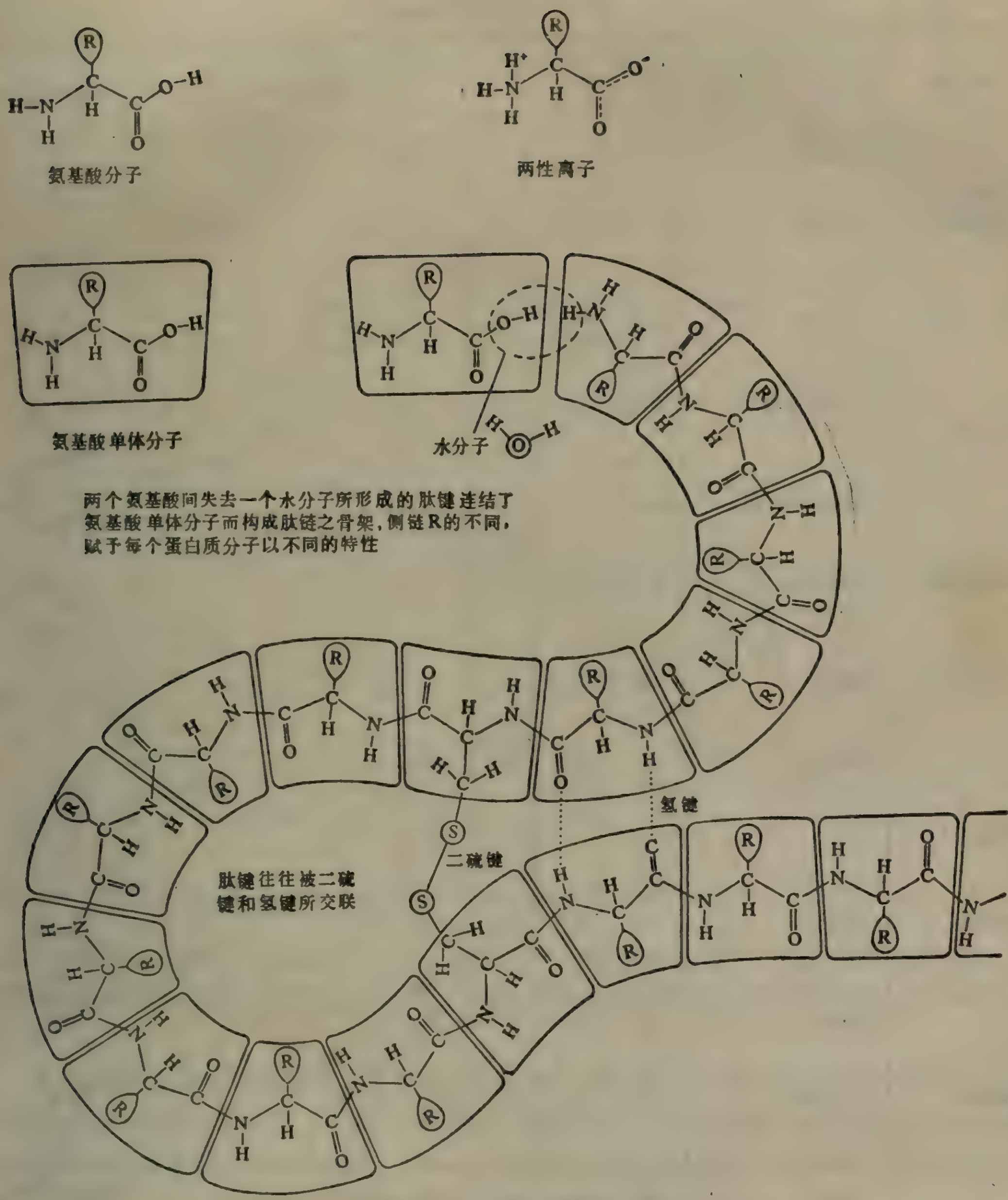
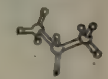
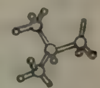
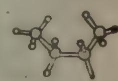
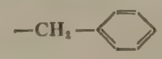
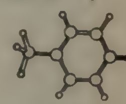
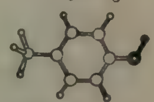
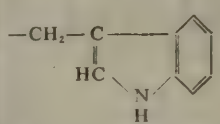
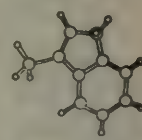
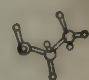
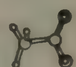
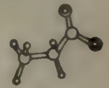
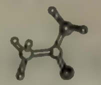
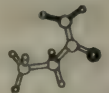


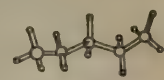
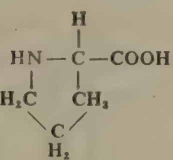
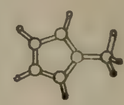
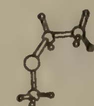
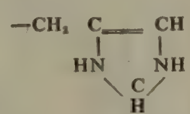
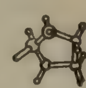
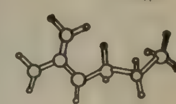
图 5-2　蛋白质分子肽链图

往由上千个氨基酸组成，为了更简洁地表达氨基酸，人们又设计了单字符，它们均列在表 5-1 中。

氨基酸的晶体结构数据和量子化学理论指数，特别是氨基酸分子中电荷分布状态是进行蛋白质结构-功能研究的基本数据之一。表 5-2 列出了氨基酸的晶体结构数据和由量子化学从头计算方法算得的氨基酸中各原子所携净电荷数。表中 MOVS，即 molecular orbital valency state，指的是原子在分子中的能量与它在孤立基态时能量的差值。原子所属类(class)是为定量

表 5-1　蛋白质的氨基酸

分　类	全　名	常用符号 中文	常用符号 英文	单字符号	R的化学结构	R的立体结构
疏水侧链						
脂肪族侧链	甘氨酸	甘	Gly	G	$-H$	
	丙氨酸	丙	Ala	A	$-CH_3$	
	缬氨酸	缬	Val	V	$-CH_3\begin{smallmatrix}CH_3\\CH_3\end{smallmatrix}$	
	亮(白)氨酸	亮(白)	Leu	L	$-CH_2-CH\begin{smallmatrix}CH_3\\CH_3\end{smallmatrix}$	
	异亮(白)氨酸	异亮(白)	Ile	I	$-CH\begin{smallmatrix}CH_3\\CH_2-CH_3\end{smallmatrix}$	
芳香环侧链	苯丙氨酸	苯(丙)	Phe	F		
	酪氨酸	酪	Tyr	Y		
	色氨酸	色	Trp	W		
亲水侧链						
羟基侧链	丝氨酸	丝	Ser	S	$-CH_2-OH$	
	苏氨酸	苏	Thr	T	$-CH(OH)-CH_3$	
羧基侧链	天冬氨酸	天冬	Asp	D	$-CH_2-C(=O)-OH$	
	谷氨酸	谷	Glu	E	$-CH_2-CH_2-C(=O)-CH$	
酰胺侧链	天冬酰胺	天酰	Asn	N	$-CH_2-C(=O)-NH_2$	
	谷氨酰胺	谷酰	Gln	Q	$-CH_2-CH_2-C(=O)-NH_2$	
	天冬氨酸或酰胺未定		Asx	B		
	谷氨酸或酰胺未定		Glx	Z		

分　类	全　　名	常用符号		单字符号	R的化学结构	R的立体结构
		中　文	英文			
亚氨基酸	精氨酸	精	Arg	R	$-CH_2-CH_2-CH_2-NH-\overset{\overset{NH}{\parallel}}{C}-NH_2$	
	脯氨酸	脯	Pro	P	全结构 	
含硫侧链	半胱氨酸	半胱	Cys	C	$-CH_2-SH$	
	甲硫氨酸	甲硫	Met	M	$-CH_2-CH_2-S-CH_3$	
碱性侧链	组氨酸	组	His	H		
	赖氨酸	赖	Lys	K	$-CH_2-CH_2-CH_2-CH_2-NH_2$	

考察分子间相互作用而引进的概念。同一类原子，意味着这些原子有相似的电子微环境。通常将具有相同的原子序数、相同的杂化状态、相似的净电荷和相似的 MOVS 的原子归划为同一类原子。

我们知道，生物分子通常是在一定的溶液条件下才能显示其功能，因此研究生物分子水溶液的结构跟研究生物分子的电子结构一样，有助于理解其结构-功能关系。氨基酸溶液结构的研究是肽和蛋白质溶液结构研究的基础。Romano 等对甘氨酸、丝氨酸、天冬氨酸、精氨酸等水溶液做了 Monte Carlo 模拟的工作，得到了一系列可与实验相比较的结果。

模拟研究的关键是水-水相互作用势和水-氨基酸相互作用势的确定。

一般，水-水相互作用势取 Matsuka 等由组态相互作用计算结果拟合得到的函数，其具体形式如下：

$$
\begin{aligned}
U_{水-水} = & q^2[1/r_{13}+1/r_{14}+1/r_{23}+1/r_{24}]+4q^2/r_{78} \\
& -2q^2[1/r_{18}+1/r_{28}+1/r_{37}+1/r_{47}]+a_1\exp[-b_1\cdot r_{56}] \\
& +a_2[\exp(-b_2\cdot r_{13})+\exp(-b_2\cdot r_{14})+\exp(-b_2\cdot r_{23})+\exp(-b_2\cdot r_{24})]
\end{aligned}
$$

表 5-2 氨基酸的结构数据*

甘 氨 酸

atom	class	x	y	z	charge	MOVS
O(2)	9	2.835	2.101	1.389	−0.50	0.56
O(1)	10	−1.320	3.200	1.631	−0.38	0.40
N	11	2.700	2.032	−3.749	−0.55	0.92
C′	5	0.673	2.824	0.416	0.46	1.41
C(A)	8	0.580	3.276	−2.431	−0.33	1.07
H(1)	1	2.597	2.270	−5.723	0.24	0.28
H(2)	1	4.433	2.698	−3.125	0.25	0.28
H(A1)	2	0.689	5.303	−2.786	0.18	0.27
H(A2)	2	−1.194	2.587	−3.213	0.22	0.30
H(O2)	4	2.607	1.920	3.160	0.40	0.32

丙 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	8.302	1.971	6.866	−0.42	0.38
O(2)	9	5.076	4.360	8.391	−0.55	0.52
N	11	7.473	3.218	2.030	−0.55	0.91
C′	5	6.389	3.299	6.575	0.50	1.36
C(A)	8	5.434	3.765	3.884	−0.13	1.11
C(B)	6	3.131	2.137	3.306	−0.59	1.01
H(A)	2	4.967	5.767	3.690	0.21	0.28
H(1)	1	8.101	1.384	2.190	0.26	0.29
H(2)	1	8.939	4.472	2.307	0.25	0.28
H(B1)	3	3.635	0.157	3.342	0.20	0.28
H(B2)	3	1.663	2.449	4.695	0.20	0.29
H(B3)	3	2.394	2.579	1.453	0.22	0.29
H(O2)	4	5.851	3.929	9.952	0.41	0.32

缬 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	15.972	10.125	8.884	−0.40	0.38
O(2)	9	14.534	6.355	7.448	−0.55	0.52
N	11	11.462	12.281	9.221	−0.50	0.91
C′	5	14.249	8.734	8.085	0.50	1.38
C(A)	8	11.653	9.897	7.716	−0.13	1.08
C(B)	7	10.963	10.369	4.916	−0.17	1.06
C(G2)	6	9.956	7.969	3.640	−0.60	1.99
C(G1)	6	13.289	11.478	3.454	−0.57	1.01
H(1)	1	11.952	11.900	11.165	0.24	0.28
H(2)	1	9.552	12.994	9.129	0.23	0.27
H(G11)	3	12.779	11.796	1.503	0.19	0.29
H(G12)	3	14.842	10.155	3.544	0.18	0.28
H(G13)	3	13.851	13.245	4.308	0.21	0.29
H(G21)	3	10.225	6.690	2.072	0.21	0.30
H(G22)	3	8.091	8.795	3.531	0.19	0.26
H(G23)	3	10.144	6.943	5.394	0.19	0.26
H(B)	2	9.477	11.768	4.870	0.19	0.26
H(A)	2	10.282	8.587	8.471	0.17	0.26
H(O2)	4	16.257	5.920	7.708	0.42	0.32

谷氨酰胺

atom	class	x	y	z	charge	MOVS
C′	5	14.907	8.505	7.153	0.49	1.39
C(A)	8	12.424	9.768	6.085	−0.12	1.12
C(B)	7	10.348	7.771	5.953	−0.38	1.03
C(G)	7	11.005	5.518	4.306	−0.42	1.01
C(D)	5	9.016	3.475	4.365	0.42	1.40
N	11	12.809	11.001	3.571	−0.55	0.91
N(E)	13	9.246	1.704	2.593	−0.63	1.02
O(1)	10	16.648	8.273	5.480	−0.40	0.36
O(2)	9	14.897	7.866	9.548	−0.54	0.53
O(E)	10	7.323	3.450	5.952	−0.42	0.37
H(1)	1	13.853	9.851	2.366	0.25	0.28
H(2)	1	11.120	11.353	2.704	0.24	0.28
H(A)	2	11.870	11.262	7.398	0.21	0.28
H(B1)	2	9.987	7.129	7.899	0.21	0.29
H(B2)	2	8.607	8.660	5.266	0.19	0.28
H(G1)	2	11.335	6.070	2.346	0.19	0.28
H(G2)	2	12.739	4.648	4.966	0.20	0.28
H(E2)	1	10.726	1.702	1.391	0.30	0.31
H(E1)	1	8.080	0.216	2.553	0.32	0.32
H(O2)	4	16.494	7.40	9.929	0.42	0.33

谷氨酸

atom	class	x	y	z	charge	MOVS
O(2)	9	4.253	10.307	23.389	−0.53	0.52
O(1)	10	2.042	6 744	23.932	−0.41	0.39
O(E1)	10	4.083	9.984	13.747	−0.41	0.38
O(E2)	9	6.751	7.406	11.764	−0.54	0.54
N	11	5.450	3.981	21.250	−0.53	0.90
C′	5	3.763	7.943	22.970	0.52	1.41
C(A)	8	5.539	6.769	21.000	−0.12	1.10
C(B)	7	4.679	7.687	18.379	−0.39	1.03
C(G)	7	6.443	7.054	16.186	−0.42	1.02
C(D)	5	5.630	8.264	13.810	0.52	1.44
H(O2)	4	2.938	11.089	24.429	0.39	0.28
H(E2)	4	6.279	8.279	10.197	0.40	0.28
H(1)	1	5.651	3.476	23.129	0.25	0.28
H(2)	1	6.992	6.195	20.258	0.23	0.27
H(A)	2	7.452	7.388	21.399	0.18	0.27
H(B1)	2	4.453	9.716	18.465	0.22	0.28
H(B2)	2	2.855	6.914	18.003	0.22	0.29
H(G1)	2	6.357	7.709	16.668	0.21	0.28
H(G2)	2	6.633	4.980	15.871	0.22	0.29

天冬氨酸

atom	class	x	y	z	charge	MOVS
O(1)	10	7.577	−0.524	11.055	−0.41	0.37
O(2)	9	9.104	−2.528	7.547	−0.54	0.52
O(D1)	10	−0.522	−2.436	6.969	−0.42	0.38
O(D2)	9	−1.277	−4.754	10.387	−0.55	0.54
N	11	4.615	−2.827	5.277	−0.65	0.86
C′	5	7.298	−1.616	8.980	0.52	1.41
C(A)	8	4.583	−1.854	7.972	−0.16	1.07
C(B)	7	2.939	−3.474	9.678	−0.57	0.88
C(G)	5	0.214	−3.512	8.866	0.55	1.55
H(1)	1	3.004	−2.887	4.708	0.32	0.30
H(2)	1	5.350	−4.444	5.207	0.29	0.29
H(A)	2	3.903	−0.121	7.868	0.26	0.30
H(B1)	2	3.013	−2.850	11.361	0.29	0.31
H(B2)	2	3.531	−5.196	9.713	0.26	0.30
H(D2)	4	−3.016	−4.687	9.898	0.42	0.33
H(O2)	4	10.669	−2.212	8.369	0.42	0.32

精 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	7.091	9.353	21.244	−0.41	0.38
O(2)	9	7.398	13.432	20.046	−0.55	0.52
N	11	11.568	7.995	19.198	−0.58	0.87
N(E)	11	4.744	8.355	11.515	−0.45	1.06
N(H1)	11	2.102	11.819	9.426	−0.61	1.04
N(H2)	12	1.290	7.388	8.700	−0.50	0.82
C′	5	8.182	11.079	20.096	0.51	1.38
C(A)	8	10.564	10.521	18.587	−0.12	1.10
C(B)	7	9.997	10.757	15.761	−0.37	1.04
C(G)	7	7.939	8.995	14.792	−0.41	1.02
C(D)	7	6.570	10.108	12.596	−0.28	1.09
C(Z)	5	2.701	9.143	9.823	0.34	1.43
H(A)	2	11.926	12.000	19.090	0.18	0.27
H(1)	1	11.454	7.718	21.081	0.27	0.28
H(2)	1	13.402	7.978	18.810	0.26	0.28
H(B1)	2	9.539	12.727	15.397	0.19	0.27
H(B2)	2	11.743	10.425	14.772	0.20	0.28
H(G1)	2	8.762	7.230	14.212	0.21	0.28
H(G2)	2	6.654	8.506	16.208	0.22	0.29
H(D1)	2	5.567	11.728	13.305	0.19	0.27
H(D2)	2	7.808	10.841	11.226	0.19	0.27
H(E)	1	5.215	6.481	11.429	0.31	0.31
H(H2)	1	3.111	13.239	10.196	0.28	0.31
H(H12)	1	0.608	12.290	8.297	0.28	0.31
H(H11)	1	−0.023	7.950	7.423	0.22	0.29
H(O2)	4	5.896	13.529	21.024	0.42	0.32

赖 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	1.457	3.059	−3.412	−0.39	0.38
O(2)	9	5.084	1.476	−5.108	−0.54	0.55
N	11	2.778	2.150	1.345	−0.55	0.91
N(Z)	11	3.923	−10.137	4.074	−0.56	0.91
C′	5	3.507	1.925	−3.248	0.50	1.36
C(A)	8	4.265	0.858	−0.659	−0.15	1.08
C(B)	7	3.826	−1.990	−0.638	−0.37	1.05
C(G)	7	4.400	−3.255	1.897	−0.36	1.06
C(D)	7	4.100	−6.110	1.736	−0.38	1.03
C(E)	8	4.260	−7.360	4.305	−0.28	1.09
H(1)	1	2.880	4.083	1.069	0.26	0.28
H(2)	1	0.918	1.591	1.315	0.26	0.29
H(A)	2	6.256	1.259	−0.325	0.18	0.27
H(B1)	2	1.863	−2.388	−1.136	0.19	0.28
H(B2)	2	5.070	−2.791	−2.089	0.17	0.27
H(G1)	2	6.333	−2.788	2.514	0.18	0.28
H(G2)	2	3.145	−2.557	3.369	0.20	0.28
H(D1)	2	2.273	−6.566	0.911	0.18	0.27
H(D2)	2	5.567	−6.889	0.506	0.17	0.27
H(E1)	2	6.067	−7.012	5.198	0.19	0.28
H(E2)	2	2.739	−6.685	5.519	0.19	0.28
H(Z1)	1	2.335	−10.525	3.068	0.25	0.28
H(Z2)	1	5.499	−10.913	3.262	0.24	0.28
H(O2)	4	6.486	0.580	−4.434	0.41	0.31

亮氨酸						
atom	class	x	y	z	charge	MOVS
O(1)	10	10.418	5.759	−1.406	−0.41	0.38
O(2)	9	13.771	7.422	0.510	−0.53	0.57
N	11	7.538	9.774	−0.638	−0.57	0.90
C′	5	11.430	7.392	−0.101	0.53	1.44
C(A)	8	9.911	9.583	0.872	−0.17	1.05
C(B)	7	9.175	9.376	3.666	−0.47	0.94
C(G)	7	11.308	9.443	5.589	−0.21	1.05
C(D1)	6	10.187	9.039	8.217	−0.69	0.92
C(D2)	6	12.710	11.956	5.490	−0.68	0.92
H(A)	2	10.905	11.090	0.648	0.21	0.27
H(B1)	2	8.044	10.761	4.042	0.23	0.29
H(B2)	2	8.291	7.782	3.896	0.25	0.29
H(G)	2	12.497	8.106	5.222	0.23	0.29
H(D11)	3	9.240	7.417	8.277	0.23	0.30
H(D12)	3	11.526	9.118	9.524	0.23	0.30
H(D13)	3	8.911	10.427	8.654	0.22	0.29
H(D21)	3	13.397	12.269	3.766	0.23	0.29
H(D22)	3	14.059	12.019	6.794	0.23	0.30
H(D23)	3	11.524	13.445	5.870	0.22	0.29
H(1)	1	6.493	11.213	−0.041	0.26	0.29
H(2)	1	7.929	9.913	−2.473	0.27	0.29
H(O2)	4	14.950	5.820	0.255	0.39	0.33

异亮氨酸						
atom	class	x	y	z	charge	MOVS
O(2)	9	11.414	5.430	9.769	−0.54	0.53
O(1)	10	12.923	9.392	8.907	−0.41	0.37
N	11	16.254	3.561	8.916	−0.50	0.91
C′	5	13.103	6.999	8.857	0.50	1.38
C(A)	8	15.369	5.839	7.445	−0.12	1.10
C(B)	7	14.634	5.076	4.741	−0.18	1.08
C(G1)	7	16.932	3.888	3.294	−0.37	1.06
C(D)	6	16.344	3.200	0.539	−0.58	1.02
C(G2)	6	13.594	7.367	3.315	−0.58	1.00
H(1)	1	16.764	4.129	10.809	0.23	0.27
H(2)	1	17.880	2.744	7.990	0.23	0.27
H(D1)	3	18.006	2.397	−0.332	0.20	0.29
H(D2)	3	15.788	4.880	−0.478	0.19	0.28
H(D3)	3	14.824	1.840	0.487	0.20	0.29
H(G21)	3	13.083	6.813	1.419	0.19	0.28
H(G22)	3	15.020	8.825	3.237	0.19	0.28
H(G23)	3	11.944	8.081	4.284	0.20	0.28
H(G11)	2	17.511	2.195	4.276	0.18	0.27
H(G12)	2	18.472	5.227	3.314	0.18	0.27
H(B)	2	13.156	3.672	4.854	0.20	0.28
H(A)	2	16.879	7.210	7.363	0.18	0.27
H(O2)	4	10.071	6.405	10.453	0.42	0.32

天冬酰胺

atom	class	x	y	z	charge	MOVS
O(2)	9	−0.027	0.889	0.007	−0.56	0.51
O(1)	10	−1.257	−3.223	−0.149	−0.39	0.38
O(D)	10	1.208	−1.967	5.312	−0.42	0.36
N(D)	13	1.128	1.187	8.217	−0.63	1.02
N	11	−4.388	−2.893	3.830	−0.54	0.92
C′	5	−1.313	−1.075	0.802	0.52	1.42
C(A)	8	−3.221	−0.540	2.889	−0.11	1.20
C(G)	5	0.174	−0.035	6.190	0.41	1.39
C(B)	7	−2.219	1.042	5.071	−0.43	0.99
H(1)	1	−3.277	−3.746	5.212	0.26	0.28
H(1)	1	−6.172	−2.540	4.576	0.24	0.28
H(D1)	1	2.643	0.512	9.079	0.33	0.32
H(D2)	1	0.301	2.785	8.801	0.31	0.31
H(A)	2	−4.741	0.643	2.055	0.20	0.28
H(B1)	2	−1.767	2.959	4.327	0.21	0.28
H(B2)	2	−3.571	1.391	6.538	0.20	0.28
H(O2)	4	0.942	0.372	−1.413	0.41	0.32

色 氨 酸

atom	class	x	y	z	charge	MOVS
O(2)	9	19.563	0.205	8.010	−0.55	0.53
O(1)	10	6.401	−0.673	10.644	−0.41	0.38
N(E)	15	−0.066	3.270	0.042	−0.46	1.13
N	11	3.129	2.855	8.055	−0.54	0.51
C(Z2)	18	5.463	−0.005	−1.321	−0.20	1.15
C(H)	18	5.445	0.308	−3.968	−0.25	1.11
C(Z3)	18	3.536	1.511	−5.302	−0.22	1.11
C(Z1)	18	1.555	2.650	−4.114	−0.23	1.11
C(E)	20	1.569	2.376	−1.487	0.14	1.12
C(D2)	19	3.456	1.079	−0.070	−0.05	1.10
C(G)	14	2.819	1.199	2.492	−0.04	1.15
C(D1)	14	0.622	2.457	2.569	−0.09	1.10
C(B)	7	4.365	−0.025	4.719	−0.43	1.04
C(A)	8	5.281	1.765	6.792	−0.15	1.08
C′	5	7.188	0.448	8.682	0.51	1.40
H(Z2)	16	6.859	−0.952	−0.311	0.21	0.31
H(H)	16	6.837	−0.437	−5.084	0.21	0.30
H(Z3)	16	3.474	1.639	−7.262	0.21	0.30
H(Z1)	16	0.000	3.606	−5.139	0.19	0.30
H(D1)	16	−0.443	3.184	4.036	0.21	0.31
H(E)	16	−1.615	4.168	−0.292	0.37	0.34
H(B1)	2	5.974	−0.827	4.081	0.23	0.29
H(B2)	2	3.164	−1.342	5.669	0.20	0.28
H(A)	2	6.084	3.387	6.042	0.26	0.30
H(1)	1	1.482	1.936	7.796	0.24	0.22
H(2)	1	2.002	1.369	8.884	0.23	0.20
H(O2)	4	10.366	−0.862	9.210	0.42	0.32

酪 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	2.183	13.293	6.896	−0.42	0.38
O(2)	9	5.659	11.372	8.565	−0.55	0.52
O(E)	9	6.280	23.914	5.176	−0.51	0.54
C′	5	4.237	12.175	6.700	0.50	1.36
C(A)	8	5.222	11.701	4.024	−0.12	1.11
C(B)	7	7.511	13.396	3.438	−0.38	1.06
C(G)	19	7.070	16.180	3.899	−0.04	1.15
C(D1)	18	5.994	17.731	2.050	−0.19	1.14
C(D2)	18	7.799	17.289	6.175	−0.19	1.13
C(E1)	18	5.665	20.311	2.454	−0.26	1.10
C(E2)	18	7.500	19.867	6.602	−0.24	1.11
C(Z)	17	6.460	21.377	4.726	0.21	1.29
N	11	3.158	12.135	2.164	−0.56	0.90
H(1)	1	2.425	13.901	2.304	0.27	0.29
H(2)	1	1.684	10.879	2.452	0.25	0.28
H(A)	2	5.796	9.722	3.851	0.21	0.28
H(B1)	2	9.066	12.752	4.621	0.20	0.28
H(B2)	2	8.082	13.070	1.477	0.21	0.29
H(D1)	16	5.458	16.934	0.237	0.21	0.30
H(D2)	16	8.666	16.127	7.635	0.20	0.30
H(E1)	16	4.837	21.484	0.994	0.18	0.30
H(E2)	16	8.120	20.721	8.362	0.22	0.31
H(E)	4	5.583	24.805	3.707	0.39	0.31
H(O2)	4	4.837	11.798	10.104	0.41	0.32

苏 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	14.342	3.170	4.536	−0.42	0.38
O(2)	9	12.936	3.109	0.464	−0.55	0.51
O(G)	9	6.086	2.147	5.125	−0.41	0.42
N	11	10.626	0.203	6.234	−0.49	0.90
C′	5	12.748	2.685	2.900	0.50	1.35
C(A)	8	10.263	1.571	3.797	−0.15	1.07
C(B)	17	8.174	3.562	4.188	0.00	1.18
C(G)	6	7.472	4.893	1.773	−0.60	0.97
H(A)	2	9.954	0.102	2.526	0.23	0.29
H(B)	2	8.540	5.162	5.393	0.18	0.27
H(G1)	3	6.006	6.263	2.152	0.18	0.28
H(G2)	3	6.793	3.527	0.417	0.21	0.29
H(G3)	3	9.113	5.843	1.018	0.20	0.28
H(1)	1	11.320	1.502	7.645	0.24	0.28
H(2)	1	11.976	−1.304	5.973	0.23	0.27
H(G)	4	3.987	3.305	5.199	0.24	0.26
H(O2)	4	14.535	3.867	0.161	0.41	0.32

丝 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	3.189	6.944	−1.571	−0.41	0.38
O(2)	9	6.521	8.809	0.419	−0.55	0.52
O(G)	9	8.385	2.913	−0.373	−0.52	0.45
N	11	2.961	2.935	1.358	−0.54	0.89
C′	5	4.853	7.003	0.099	0.52	1.41
C(A)	8	4.953	4.814	1.988	−0.14	1.08
C(B)	17	7.556	3.614	2.068	−0.18	1.13
H(A)	2	4.499	5.562	3.875	0.17	0.26
H(2)	1	1.245	3.878	1.314	0.26	0.28
H(1)	1	2.881	1.591	2.791	0.23	0.27
H(B1)	2	8.887	5.012	2.811	0.21	0.29
H(B2)	2	7.558	1.993	3.354	0.17	0.27
H(G)	4	7.790	1.198	−0.754	0.36	0.30
H(O2)	4	6.233	10.053	−0.842	0.42	0.32

胱 氨 酸

atom	class	x	y	z	charge	MOVS
O(1)	10	6.623	2.652	4.243	−0.41	0.40
O(2)	9	4.099	4.782	6.862	−0.54	0.51
N	11	6.049	5.961	0.467	−0.54	0.90
C′	5	5.326	4.448	4.744	0.53	1.42
C(A)	8	4.946	6.659	2.950	−0.13	1.09
C(B)	21	2.191	7.437	2.730	−0.50	0.97
S(G)	22	0.008	4.954	1.926	0.11	0.19
H(2)	1	5.698	7.391	−0.849	0.24	0.28
H(1)	1	7.970	5.727	0.535	0.26	0.29
H(A)	2	6.048	8.245	3.703	0.18	0.27
H(B1)	2	1.560	8.184	4.567	0.22	0.29
H(B2)	2	1.993	9.013	1.369	0.19	0.28
H(O2)	4	4.421	3.369	8.062	0.40	0.32
O(O1)	10	−6.623	2.652	−4.243	−0.41	0.40
O(O2)	9	−4.099	4.782	−6.862	−0.54	0.51
N(N)	11	−6.049	5.961	−0.467	−0.54	0.90
C(C′)	5	−5.326	4.448	−4.744	0.53	1.42
C(CA)	8	−4.946	6.659	−2.950	−0.13	1.09
C(CB)	21	−2.191	7.437	−2.730	−0.50	0.97
S(SG)	22	−0.008	4.954	−1.926	0.11	0.19
H(H2)	1	−5.698	7.391	0.849	0.24	0.28
H(H1)	1	−7.970	5.727	−0.535	0.26	0.29
H(HA)	2	−6.048	8.245	−3.703	0.18	0.27
H(HB1)	2	−1.560	8.184	−4.567	0.22	0.29
H(HB2)	2	−1.993	9.013	−1 369	0.19	0.28
H(HO2)	4	−4.421	3.369	−8.062	0.40	0.32

脯 氨 酸

atom	class	x	y	z	charge	MOVS
C′	5	10.362	12.748	2.204	0.52	1.41
C(A)	8	9.759	10.011	2.835	−0.13	1.06
C(B)	7	9.521	9.528	5.700	−0.38	1.03
C(G)	7	7.942	7.161	5.710	−0.40	1.01
C(D)	7	6.493	6.989	3.272	−0.30	1.06
N	11	7.353	9.214	1.683	−0.39	0.90
O(1)	10	9.175	13.964	0.594	−0.41	0.38
O(2)	9	12 209	13.670	3.578	−0.55	0.51
H(1)	1	7.631	8.652	−0 259	0.25	0.28
H(D1)	2	4.486	7.100	3.625	0.20	0.28
H(D2)	2	6.906	5.232	2.320	0.18	0.27
H(G1)	2	6.640	7.211	7.281	0.21	0.28
H(G2)	2	9.153	5.530	5.904	0.20	0.28
H(B1)	2	11.342	9.225	6.569	0.21	0.29
H(B2)	2	8.579	11.069	6.651	0.20	0.29
H(A)	2	11.253	8.831	2.099	0.18	0.27
H(O2)	4	12.439	15.386	3.106	0.41	0.32

羟脯氨酸

atom	class	x	y	z	charge	MOVS
O(2)	9	9.460	8.278	21.723	−0.54	0.52
O(1)	10	6.828	4.966	22.454	−0.42	0.37
O(D)	9	3.818	7.431	13.469	−0.53	0.45
N	11	6.848	9.892	17.686	−0.40	0.91
C′	5	7.483	6.864	21.236	0.51	1.39
C(A)	8	5.878	7.540	18.928	−0.13	1.08
C(B)	7	6.037	5.501	16.894	−0.41	1.00
C(G)	17	6.246	6.883	14.384	0.00	1.14
C(D)	7	7.689	9.274	15.078	−0.31	1.04
H(1)	1	5.393	11.229	17.695	0.26	0.28
H(A)	2	3.952	7.895	19.525	0.21	0.28
H(B1)	2	7.657	4.349	17.190	0.20	0.28
H(B2)	2	4.417	4.293	16.927	0.23	0.30
H(G)	2	7.300	5.788	12.983	0.17	0.27
H(D1)	4	3.966	7.703	11.661	0.36	0.30
H(B1)	2	9.700	8.918	15.147	0.20	0.28
H(B2)	2	7.310	10.812	13.841	0.20	0.29
H(O2)	4	10.286	7.590	23.160	0.41	0.32

组氨酸

atom	class	x	y	z	charge	MOVS
O(1)	10	2.200	0.065	7.068	−0.41	0.38
O(2)	9	−1.836	0.142	8.618	−0.55	0.52
N	11	4.031	−2.474	11.080	−0.70	0.83
N(D)	12	5.215	−0.837	15.940	−0.29	0.83
N(E)	15	3.662	−0.007	19.734	−0.46	1.14
C′	5	0.609	−0.235	8.783	0.52	1.41
C(A)	8	1.611	−1.059	11.400	−0.15	1.08
C(B)	7	2.044	1.255	13.101	−0.52	0.95
C(G)	14	3.015	0.551	15.680	0.07	1.15
C(E)	14	5.507	−1.129	18.411	−0.03	1.09
C(D)	14	2.047	1.071	18.008	−0.15	1.07
H(1)	1	4.764	−2.574	12.588	0.34	0.30
H(2)	1	5.126	−1.702	10.020	0.31	0.30
H(E2)	16	3.414	0.000	21.503	0.39	0.35
H(A)	2	0.401	−2.242	12.196	0.24	0.29
H(B1)	2	0.430	2.103	13.230	0.25	0.29
H(B2)	2	3.131	2.546	12.303	0.24	0.29
H(E1)	16	6.887	−2.020	19.114	0.25	0.33
H(D)	16	0.519	2.020	18.436	0.23	0.33
H(O2)	4	−2.196	0.691	6.947	0.42	0.32

半胱氨酸

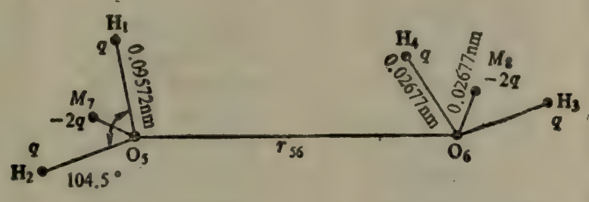
atom	class	x	y	z	charge	MOVS
S(G)	22	0.008	4.954	1.926	0.06	0.22
O(1)	10	6.623	2.652	4.243	−0.41	0.40
O(2)	9	4.099	4.782	6.862	−0.54	0.51
N	11	6.049	5.961	0.467	−0.54	0.90
C′	5	5.326	4.448	4.744	0.53	1.41
C(A)	8	4.946	6.659	2.950	−0.13	1.09
C(B)	21	2.191	7.437	2.730	−0.50	0.95
H(2)	1	5.698	7.391	−0.849	0.24	0.28
H(1)	1	7.970	5.727	0.535	0.26	0.29
H(A)	2	6.048	8.245	3.703	0.18	0.27
H(B1)	2	1.560	8.184	4.567	0.20	0.27
H(B2)	2	1.993	9.013	1.369	0.19	0.28
H(S0)	23	0.019	4.954	4.477	0.06	0.22
H(O2)	4	4.421	3.369	8.062	0.40	0.32

甲硫氨酸

atom	class	x	y	z	charge	MOVS
S	22	0.306	1.528	3.633	0.21	0.20
O(2)	9	−4.937	1.726	13.260	−0.55	0.52
O(1)	10	−1.638	−0.693	14.122	−0.43	0.38
N	11	1.782	2.718	12.282	−0.51	0.91
C′	5	−2.556	1.084	13.007	0.51	1.37
C(A)	8	−0.882	2.718	11.309	−0.13	1.11
C(B)	7	−0.918	1.501	8.542	−0.38	1.05
C(G)	21	0.684	2.984	6.676	−0.51	0.97
C(E)	21	2.718	3.109	1.904	−0.74	0.94
H(G1)	2	2.644	2.934	7.240	0.21	0.28
H(G2)	2	0.063	4.927	6.609	0.20	0.28
H(B1)	2	−2.844	1.454	7.867	0.20	0.28
H(B2)	2	−0.205	−0.409	8.652	0.20	0.28
H(A)	2	−1.604	4.626	11.246	0.21	0.29
H(1)	1	2.502	0.808	12.294	0.24	0.28
H(2)	1	1.813	3.470	14.179	0.23	0.28
H(E1)	3	3.676	4.433	3.127	0.21	0.29
H(E2)	3	1.885	4.112	0.334	0.21	0.29
H(E3)	3	4.059	1.740	1.201	0.21	0.29
H(O2)	4	−5.696	0.578	14.414	0.41	0.32

$$+a_3[\exp(-b_3\cdot r_{16})+\exp(-b_3\cdot r_{26})+\exp(-b_3\cdot r_{35})+\exp(-b_3\cdot r_{45})]$$
$$-a_4[\exp(-b_4\cdot r_{16})+\exp(-b_4\cdot r_{26})+\exp(-b_4\cdot r_{36})+\exp(-b_4\cdot r_{45})] \quad (5.1)$$

式中 q、a_1、a_2、a_3、a_4、b_1、b_2、b_3、b_4 是拟合系数，距离 r_{ij} 的定义见图 5-3。图中的 M点在水分子的 C_{2v} 对称轴上离氧 原子 0.02677 nm 处。 M点上置有 点电荷 $-2q$。 图中 下部还列出了$U_{水-水}$公式中各拟合参数值。



q^2	a_1	a_2	a_3	a_4
0.565117	1864.271482	0.662712	2.684452	0.675342
	b_1	b_2	b_3	b_4
	2.753110	1.299982	1.439787	1.141494

图 5-3　水-水相互作用势中距离的定义和参数值

氨基酸-水相互作用势采用 Clementi 等人由从头计算法计算结果拟合得到的 1-6-12 势：

$$U_{A-W}=\sum_i\sum_j[-A_{ij}^{ab}/r_{ij}^6+B_{ij}^{ab}/r_{ij}^{12}+C_{ij}^{ab}q_iq_j/r_{ij}] \quad (5.2)$$

其中 A、B、C 是拟合常数；i、j 分别表示氨基酸 和水分子中的原子；指标 a 表示氨基酸中下标为 i 的原子所属的类(class)；指标 b 表示水分 子中下标为 j 的原子所属的类（因为水分 子中只有氢、氧二类原子，所以 b 实际上是起了区分水分子中氢、氧原子的作用）；q_i、q_j 是原子 i（氨基酸中的)和原子 j（水分子中的）的净电荷。 表 5-3 a 和 b 是氨基酸-水相互 作用势的参数表（距离单位：Å，能量单位：kcal/mol)。注意，该表需与表 5-2 联合使用。

在 300 K，用蒙特卡洛方法模拟了中性和两性的丝氨酸分子的溶液结构。 模拟系统采用立方元胞，内含 250 个水分子和一个丝氨酸分子，元胞边长由实验值(密度测定)确定。计算中采用了周期性边界条件。

模拟计算表明，在图 5-4 中的 S1、S2 和 S3 这三种溶液中，水-水相互作用能为-731.34 ± 2.39、-726.56 ± 2.39 和-736.12 ± 2.39 kcal/mol，丝氨酸-水相互作用能为-4206.4 ± 71.7、-10205.3 ± 119.5 和-8460.6 ± 95.6 kcal/mol。

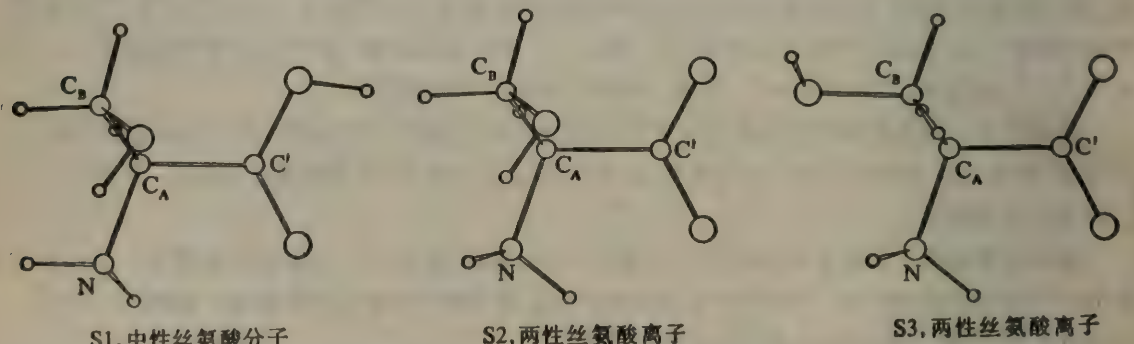
图 5-5 给出了 S1、S2 和 S3 溶液中 N 原子和 C′ 原子的径向分布函数 RDF 和偶极矩 取向关联函数 OCF。

径向分布函数，简单地讲，它相当于在一个粒子周围距离为 r 的地方出现另一个粒子的几率相对于随机分布的比值。由于随机分布的径向分布函数为 1，所以径向分布函数可以理解为在距离 r 处找到另一粒子的几率。

偶极矩取向关联函数可理解如下：求出分子偶极矩 方向的单位矢量以及从中心原子到分子中某一指定原子连线方向上的单位矢量，上述单位矢量间的内积称为偶极矩取向关联函数。

表 5-3a　氨基酸和水分子中氧原子相互作用参数

class	A	B	C
1	0.207 521 847+01	0.173 850 511+04	0.100 088 759+01
2	0.824 252 534+01	0.145 269 011+04	0.100 000 000+01
3	0.130 720 375+01	0.280 374 992+03	0.998 704 314+00
4	0.159 769 344+01	0.463 124 774+04	0.100 093 982+01
5	0.356 004 411+01	0.184 585 125+06	0.999 759 927+00
6	0.356 004 411+01	0.389 186 260+05	0.100 027 880+01
7	0.190 353 591+02	0.618 960 681+04	0.100 000 000+01
8	0.117 872 994+03	0.235 846 649+04	0.100 101 370+01
9	0.845 311 871+02	0.870 600 693+03	0.100 005 104+01
10	0.163 284 536+03	0.762 464 880+04	0.999 389 425+00
11	0.141 507 664+03	0.277 916 666+04	0.100 009 498+01
12	0.491 524 463+02	0.486 883 081+04	0.998 375 371+00
13	0.347 660 562+01	0.682 350 397+04	0.100 299 819+01
14	0.968 582 392+01	0.130 161 777+06	0.100 080 614+01
15	0.441 930 092+03	0.127 116 438+05	0.100 051 089+01
16	0.356 528 398+01	0.104 377 336+04	0.998 747 133+00
17	0.356 004 411+01	0.121 598 151+06	0.992 949 978+00
18	0.356 004 411+01	0.404 495 615+05	0.100 236 642+01
19	0.356 004 411+01	0.157 257 820+06	0.874 550 834+00
20	0.385 323 176+02	0.177 231 436+06	0.981 899 209+00
21	0.604 173 851+01	0.545 685 400+05	0.100 149 776+01
22	0.259 413 876+03	0.918 035 120+04	0.101 107 869+01
23	0.102 959 207+03	0.776 556 488+04	0.102 515 377+01



S1,中性丝氨酸分子　　S2,两性丝氨酸离子　　S3,两性丝氨酸离子

图 5-4　丝氨酸分子沿 z 轴的投影图

S2、S3是两种不同构象的两性丝氨酸离子

从能量结果中可以看到,无论是 S1、S2 还是 S3，它们的 $E_{水-水}$基本一致(由 200 个水分子包围一个甘氨酸中性分子或一个两性甘氨酸分子离子所构成的立方元胞,其 $E_{水-水}$的模拟结果也分别为-724.17 ± 2.39 和-719.39 ± 2.39 kcal/mol)。这表明模拟中所用的水分子数目已能满意反映真实体系中溶质-溶剂(这里实为氨基酸-水)相互作用。两性离子和水相互作用比中性分子大，这主要是由于两性离子氨端水化区的溶质-溶剂相互作用比中心分子大得多造成的。S 2 和

表 5-3 b　氨基酸与水分子中氢原子相互作用参数

class	A	B	C
1	0.227 048 409+03	0.319 285 464+04	0.998 202 369+00
2	0.344 879 279+01	0.332 071 268+04	0.999 966 606+00
3	0.344 879 279+01	0.704 513 690+04	0.998 410 933+00
4	0.389 675 373+03	0.307 583 102+04	0.100 093 916+01
5	0.242 698 627+03	0.570 163 312+04	0.999 941 483+00
6	0.102 073 144+02	0.294 559 020+06	0.999 955 520+00
7	0.110 689 075+03	0.276 483 570+06	0.999 073 327+00
8	0.102 073 144+02	0.3459 47 762+06	0.100 345 021+01
9	0.112 916 672+02	0.305 269 293+06	0.999 645 844+00
10	0.102 073 144+02	0.261 569 141+06	0.100 092 629+01
11	0.124 076 189+02	0.6318286125+06	0.999 096 863+00
12	0.220 215 490+02	0.200 655 234+06	0.996 603 809+00
13	0.318 320 038+03	0.615 968 625+06	0.100 005 582+01
14	0.126 181 303+03	0.130 406 014+06	0.101 622 795+01
15	0.161 787 306+04	0.626 337 617+06	0.100 060 296+01
16	0.233 495 670+03	0.792 550 488+04	0.100 239 015+01
17	0.102 073 144+02	0.102 467 840+06	0.993 570 738+00
18	0.102 073 144+02	0.520 450 148+06	0.999 982 804+00
19	0.199 325 260+03	0.136 718 561+05	0.109 999 999+01
20	0.655 663 666+03	0.911 956 562+06	0.101 060 666+01
21	0.229 295 538+03	0.155 507 061+05	0.100 122 450+01
22	0.139 064 225+02	0.174 304 908+07	0.995 094 240+00
23	0.254 667 215+03	0.369 866 099+05	0.105 506 256+01

S3 的 $E_{丝-水}$ 的差别则是由它们的疏水集团 $\rangle CH_2$ 和两个亲水性基团空间取向的不同而造成的。此外,S2 在氨端水化区域中的溶质-溶剂相互作用范围比 S3 的要更大些。

以下再简略讨论每一水化区域的结构特点。

S1 的 N/H(以 N 原子为中心,考察 H 原子的情况)RDF 在 0.3 nm 和 0.35 nm 有两个峰;N/O 的 RDF 峰在 0.3 nm 处;在 0.4—0.6 nm 存在 N/O 的 RDF 宽峰;N/DIP OCF 显示出取向有序,且一直延伸到 0.45 nm。和甘氨酸水溶液相比,丝氨酸水溶液中该区的有序性大得多。在甘氨酸溶液中与甘氨酸中 N 原子最接近的水分子取向是水分子中的 H 原子指向 N 原子, 但在S1 中却不是如此,这是因为 S1 存在—OH 羟基的缘故。

S2 的 N/H 之 RDF 在 0.3nm 和 0.35 nm 处各 有一个峰, N/O 的 RDF 在 0.28 nm 有一尖锐峰, N/DIP OCF 表现出该区已存在高度的取向有序性。

S3 的情况与 S2 类似, 所不同的是 N/H 的 RDF 峰在 0.3—0.35 nm 之间, N/O 的 RDF 峰在 0.27 nm, C′/O 的RDF 峰在 0.35 nm 处。

S1 的 C′/H RDF 在 0.35 nm有一宽峰, C′/O RDF 在 0.32nm 有一主峰, 在 0.35—0.5 nm 有一较低的宽峰, 0.5 nm 以外仍显示取向有序。

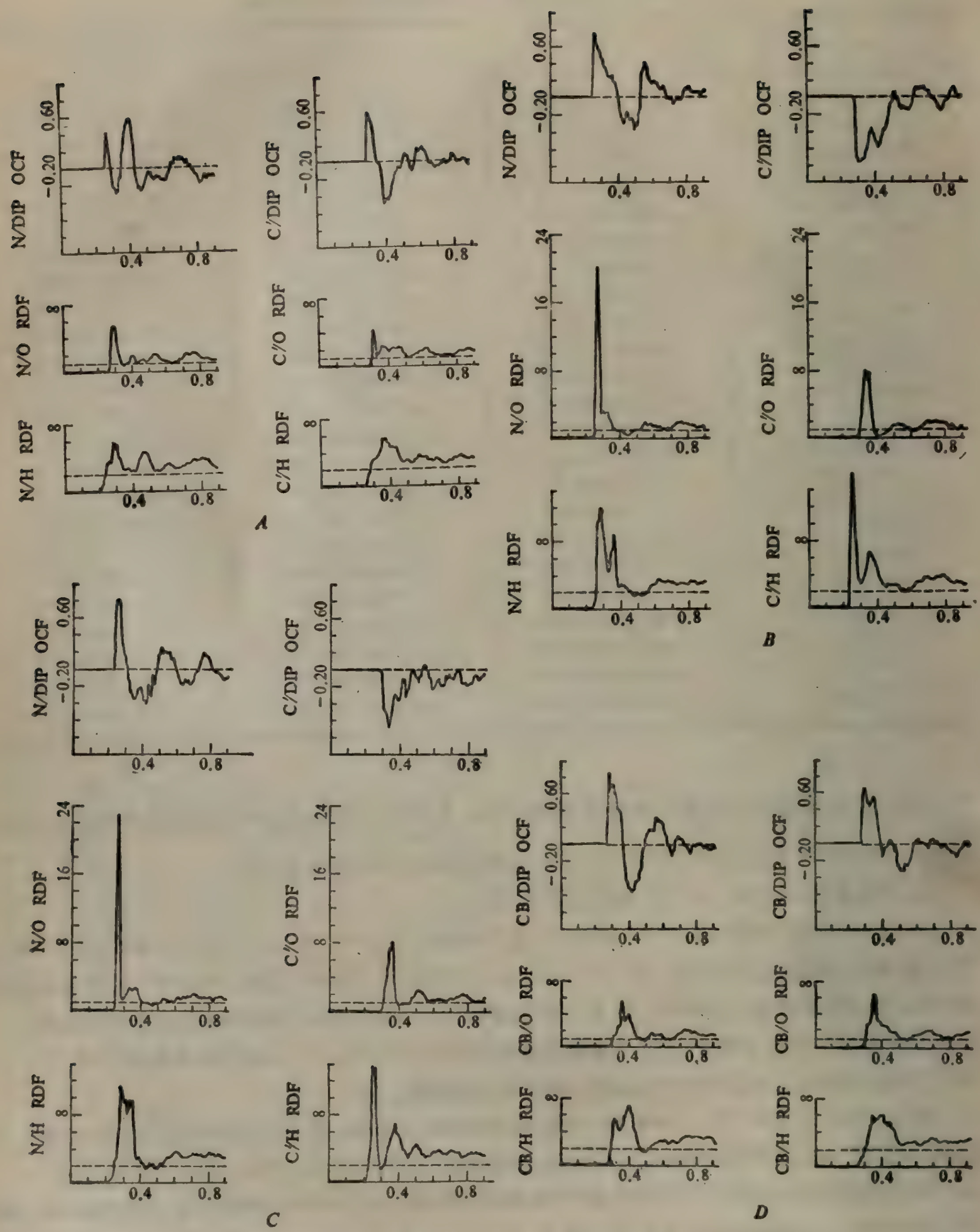


图 5-5 丝氨酸的 RDF_s 和 OCF_s[横坐标为R(nm)]

A. S1中N和C′区域的 RDF_s 和 OCF_s,

B. S2中N和C′区域的 RDF_s 和 OCF_s,

C. S3中N和C′区域的 RDF_s 和 OCF_s,

D. S1(左)和S2(右) C_β 的 RDF_s 和 OCF_s。

S 2 的 C′/H RDF 在 0.27 nm 和 0.37 nm、C′/O 的 RDF 在 0.35 nm 处均有主峰。取向有序延伸到 0.4 nm。

S 3 的情况与 S 2 类似。

丝氨酸模拟结果的总图象与甘氨酸水溶液模拟结果定性相符。

模拟结果的进一步分析还告诉我们，在丝氨酸溶液中，有若干个水分子在氨基和羧基处是直接化学水化；溶液中存在着由氢键键合的复杂水分子网络，这些网络结构单元有的就在溶质分子周围，有的直接与亲水的氨基和羧基水化；在溶质分子亲水区中往往存在“水化氛”。值得指出的是，模拟计算表明在丝氨酸溶液中不存在与—OH 羟基直接键合的强度较大的氢键，这很可能是由于疏水性 $\rangle CH_2$ 基团的存在及更为强烈的亲水性基团竞争的结果。

蒙特卡洛模拟最终给出系统进入平衡态后的结构状态，故除了上面所提到的信息外，还可以抽提出每个溶剂分子，特别是溶质分子周围溶剂分子的地位等精细微环境信息，在此基础上还能得到系统的平衡热力学性质、结构动力学性质。下面要介绍的就是这方面工作的一个典型例子。

5.2 短杆菌肽A离子选择性的蒙特卡洛研究

生物膜的离子选择性在许多生物机制中起着重要的作用。短杆菌肽 A 通道是目前为数不多的确定了空间结构的离子选择性跨膜之一。该膜对阴离子和二价阳离子是不渗透的，对单价阳离子有明显的离子选择性。若以对 Na^+ 的渗透性为标准(1)的话，那么有：

$H^+(150) > NH_4^+(8.9) > Cs^+(5.8) > Rb^+(5.5) > K^+(3.9) > Na^+(1) > Li^+(0.33)$

溶剂化短杆菌肽 A 跨膜通道离子选择性的蒙特卡洛计算的主要目的是比较 Na^+、K^+ 阳离子-水分子-短杆菌肽 A 体系的能量、能量组成，分析它们对离子选择性造成的影响，研究 Na^+、K^+ 存在时，短杆菌肽 A 通道的水化结构。

短杆菌肽 A 是由 L-氨基酸和 D-氨基酸交替组成的 C 端和 N 端均被保护的 15 肽。一级结构为 HCO—L.Val^1—Gly^2—L.Ala^3—D.Leu^4—L.Ala^5—D.Val^6—L.Val^7—D.Val^8—L.Trp^9—D.Leu^{10}—L.Trp^{11}—D.Leu^{12}—L.Trp^{13}—D.Leu^{14}—L.Trp^{15}—$NHCH_2CH_2OH$。

图 5-6 是短杆菌肽 A 膜通道。

模拟的能量结果如下：水-水相互作用能 S 1、S 2 和 S 3 分别为 -30.6 ± 0.1、-30.4 ± 0.1 和 -30.8 ± 0.1 kJ/mol。

图 5-7 是半径为 0.35 nm 时短杆菌肽 A 的圆柱形图以及它们在 R-Z 面上的投影。

从图中可以看出短杆菌肽 A 跨膜由短杆菌肽 A 双体组成，每个单体都是一个 β 螺旋结构，每 6.3 个残基构成螺旋的一圈。肽的 C—O 部分交替地指向 L 残基的羰基端和 D 残基的氨端，A、B 两个单体相对而立。

在对 Na^+-K^+-H_2O-短杆菌肽 A 系统进行蒙特卡洛模拟时，水-水、水-阳离子，阳离子-短杆菌肽 A、水-短杆菌肽 A 的相互作用势函数均由量子化学从头计算法得出。模拟温度为300K，模拟体积为半径 0.55 nm、长度 4.8 nm 的圆柱体。圆柱体内包含一个短杆菌肽 A 通道，一个阳离子和

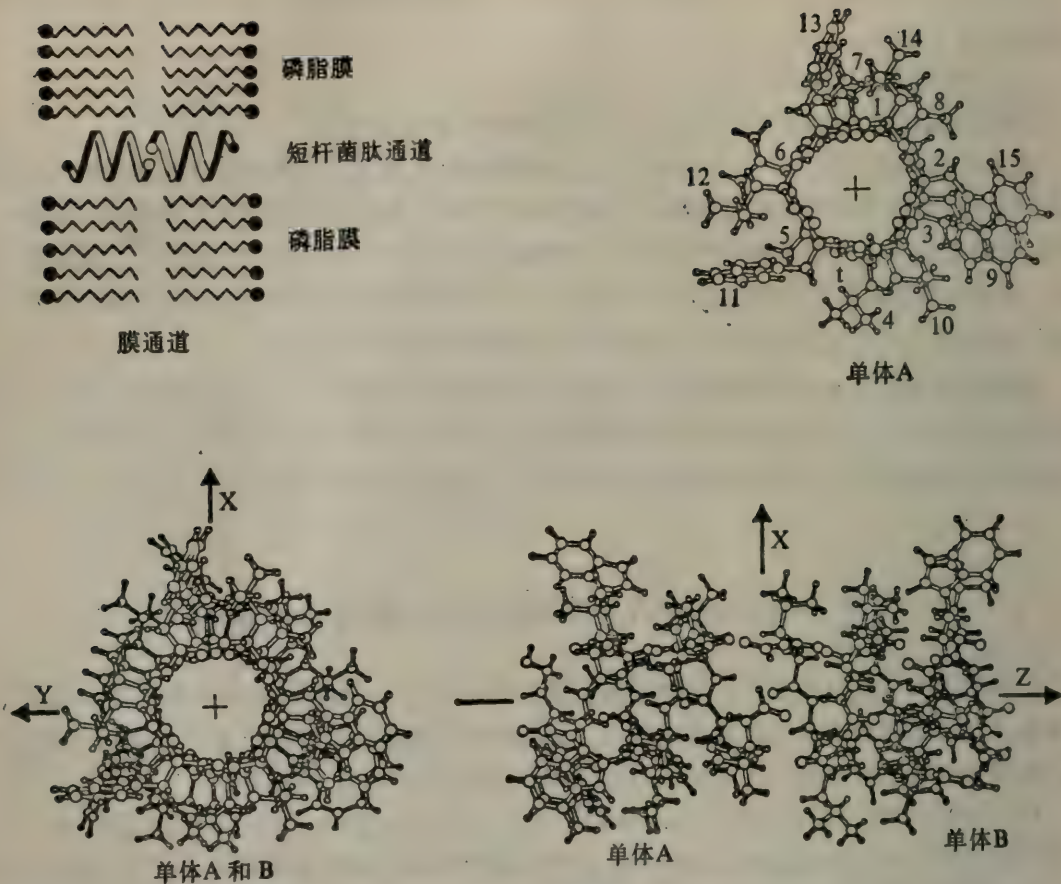


图 5-6 短杆菌肽A膜通道的特征

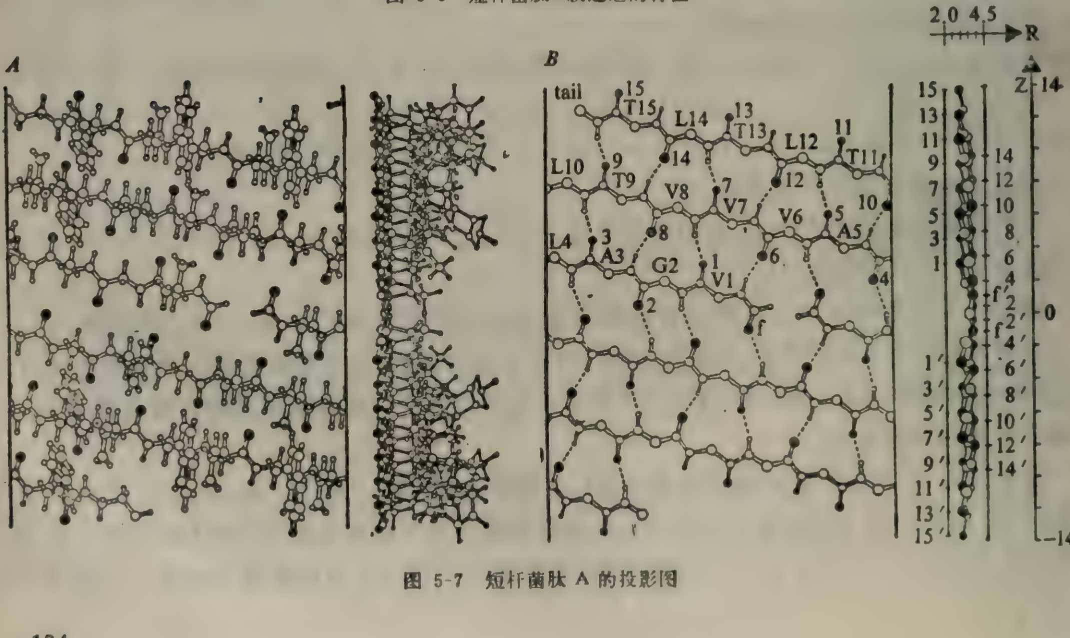


图 5-7 短杆菌肽 A 的投影图



图 5-8　$z=0,0.2,0.4,0.6,\cdots\cdots,1.4$ nm 时，等能面在 X-Y 面上的投影

左图表示 $z-0.3$ nm 到 $z+0.3$ nm 范围内短杆菌肽A的组成部分。第二、三、四列分别是 H_2O、K^+和 Na^+ 与短杆菌肽A相互作用的等能图，其中每幅图左下方的数值表示等能图中最小能量值(kcal/mol)。

第一列的范围为 $x=\pm0.6$ nm，$y=\pm0.6$ nm，其余的皆为 $x=\pm6$ 原子单位，$y=\pm6$ 原子单位。

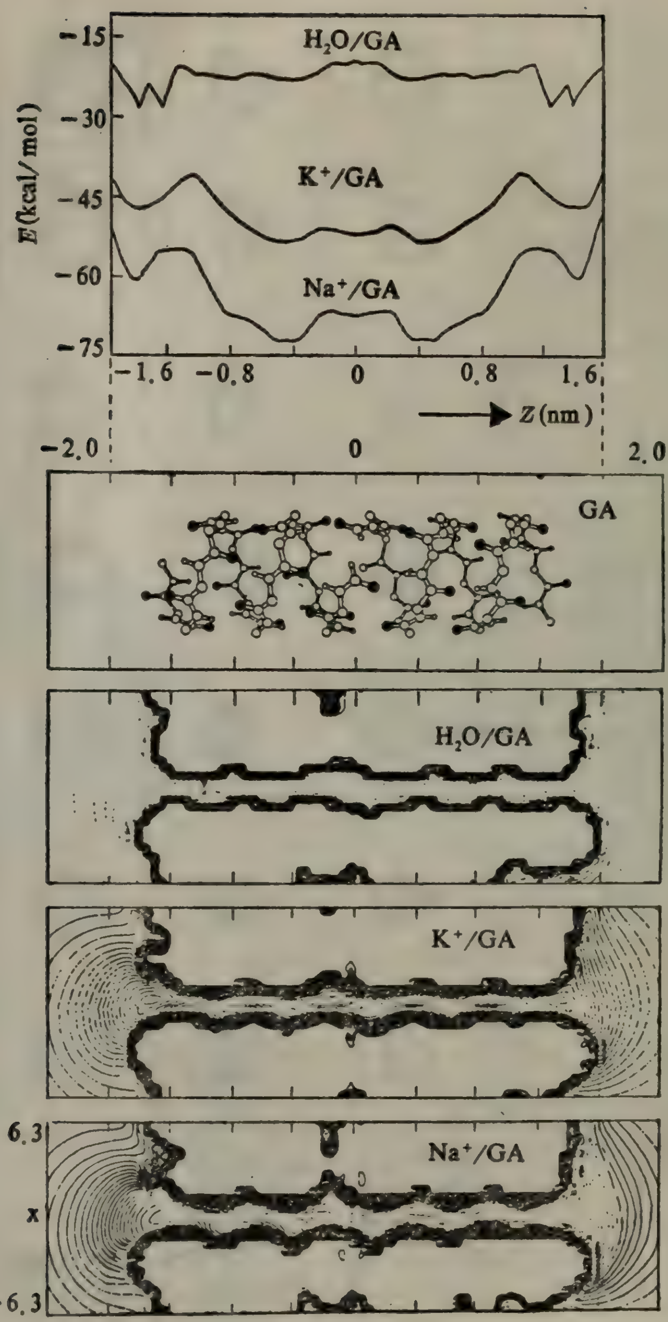


图 5-9 在短杆菌肽A通道内 H_2O、K^+ 和 Na^+ 与短杆菌肽A相互作用能的极小值与 z 的关系及其在 X-Z 平面上的投影

81 个水分子。在上述条件下，水的密度为 1 。水分子和阳离子可在圆柱体内自由运动，对水分子加了 z 方向平移对称性条件。每一系统计算了一百万步，其中的后五十万步用于统计分析。

从图 5-7 中可以看到，羰基氧处在短杆菌肽 A 螺旋两条相邻的螺线之间，但离氧与之键连的那条螺线更近。可以想象，阳离子的位置也应如此。另外，构成螺旋的原子会对离子产生近程排斥作用，因此阳离子也必落在两条螺旋之间。

图 5-8 至图 5-10 是 Na^+、K^+-H_2O-短杆菌肽 A 系统能量的各种关系示意。

从图中可以清楚地看出，Na^+/(短杆菌肽)比 K^+/(短杆菌肽)要大，这是由于Na^+与短杆菌肽 A 的原子的相互作用比 K^+ 的要强。由此可以推断，在短杆菌肽 A 中，Na^+ 比 K^+ 要更容易进入能量适宜的位置，也就是说，K^+偏离 z 轴运动的倾向比 Na^+要小。

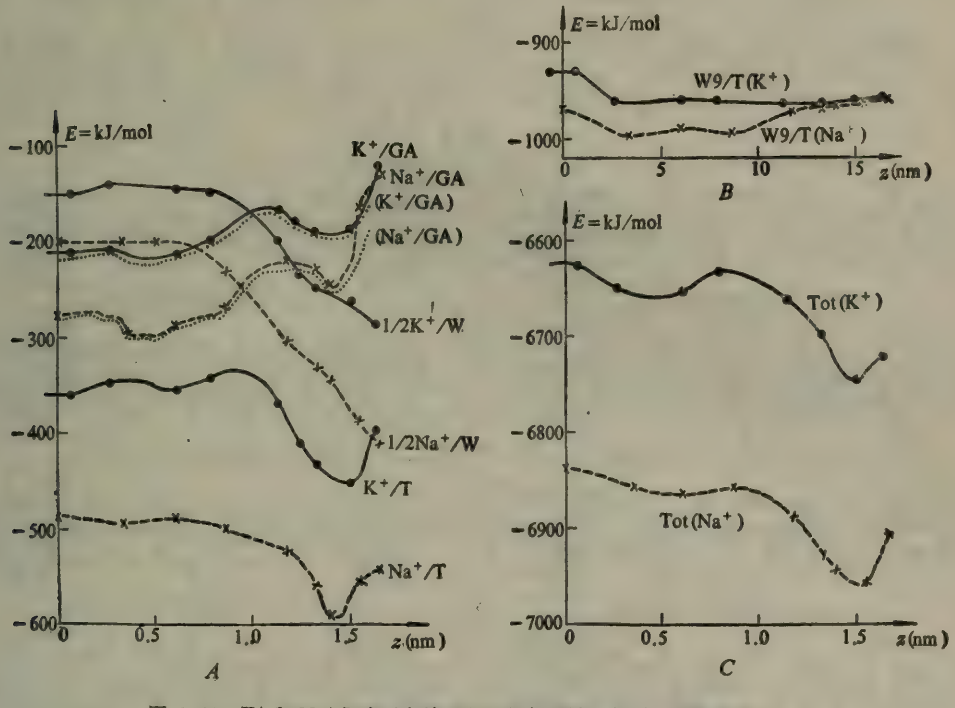


图 5-10 K⁺ 和 Na⁺ 阳离子与溶剂化的短杆菌肽 A 相互作用能曲线。
A 表示阳离子贡献部分，*B* 表示通道中 9 个水分子的贡献，*C* 表示 K⁺ 和 Na⁺ 系统的总能量。

Na^+/(短杆菌肽 A)、K^+/(短杆菌肽 A)所构成的势能、羰基氧所造成的深势阱、链上原子对阳离子所必然产生的近程排斥作用以及短杆菌肽 A 的结构特征，决定了阳离子 Na^+、K^+ 在膜通道中的运动途径，如图 5-11 所示。

图 5-12 和图 5-13 是 Na^+、K^+ 存在时，短杆菌肽 A 通道中水分子的键合结构状态。

分析上述两图可以看出，无论是 Na^+ 还是 K^+，在通道内(−1.2 nm 至 1.2 nm) 存在 9 个水分子。渗透和扩散实验值为 6—7 个水分子，电极动力学实验值为 12 个水分子。理论计算值介于两类实验值之间。

通道中的水成链状束。通道中的阳离子在每一水化层中被 2 个水分子溶剂化。当阳离子运

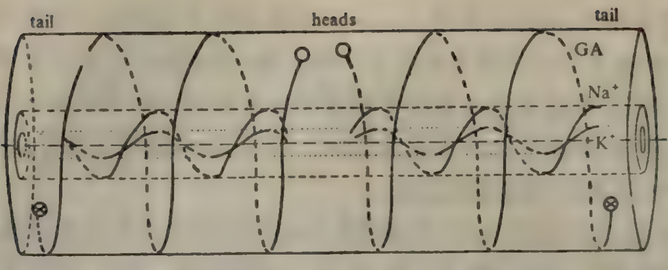


图 5-11 在短杆菌肽A通道内 K^+ 和 Na^+ 的运动途径。
图中也画出了短杆菌肽A骨架的螺旋型模式。

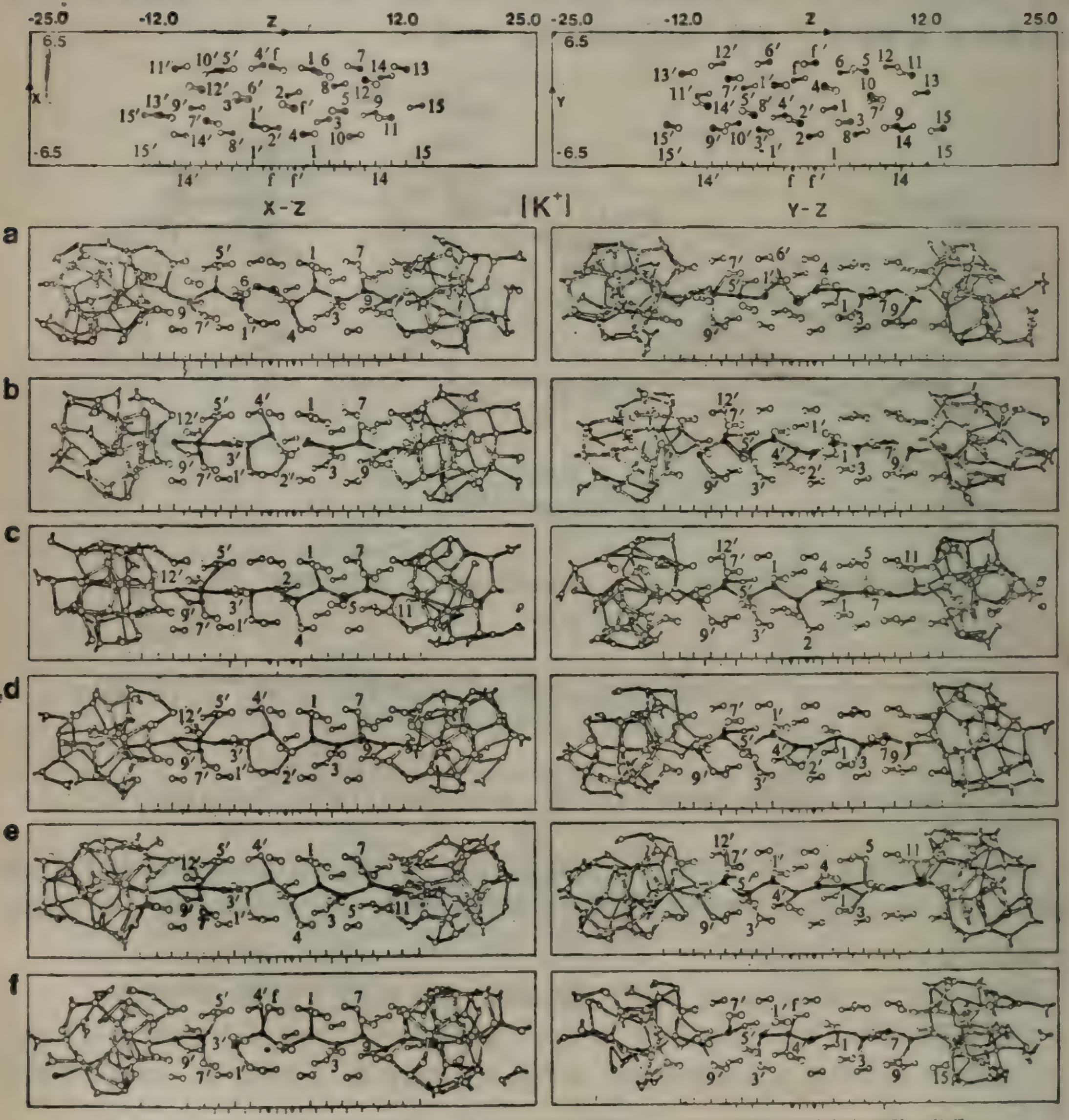


图 5-12 K^+存在时，溶剂化短杆菌肽A通道内的键合结构图。顶部两张图只画出短杆菌肽A的骨架基团。其余的图表示K^+的平均位置处于 a: $z=-0.08$, b: $z=0.27$, c: $z=0.61$, d: $z=0.79$, e: $z=1.14$, f: $z=1.50$ nm

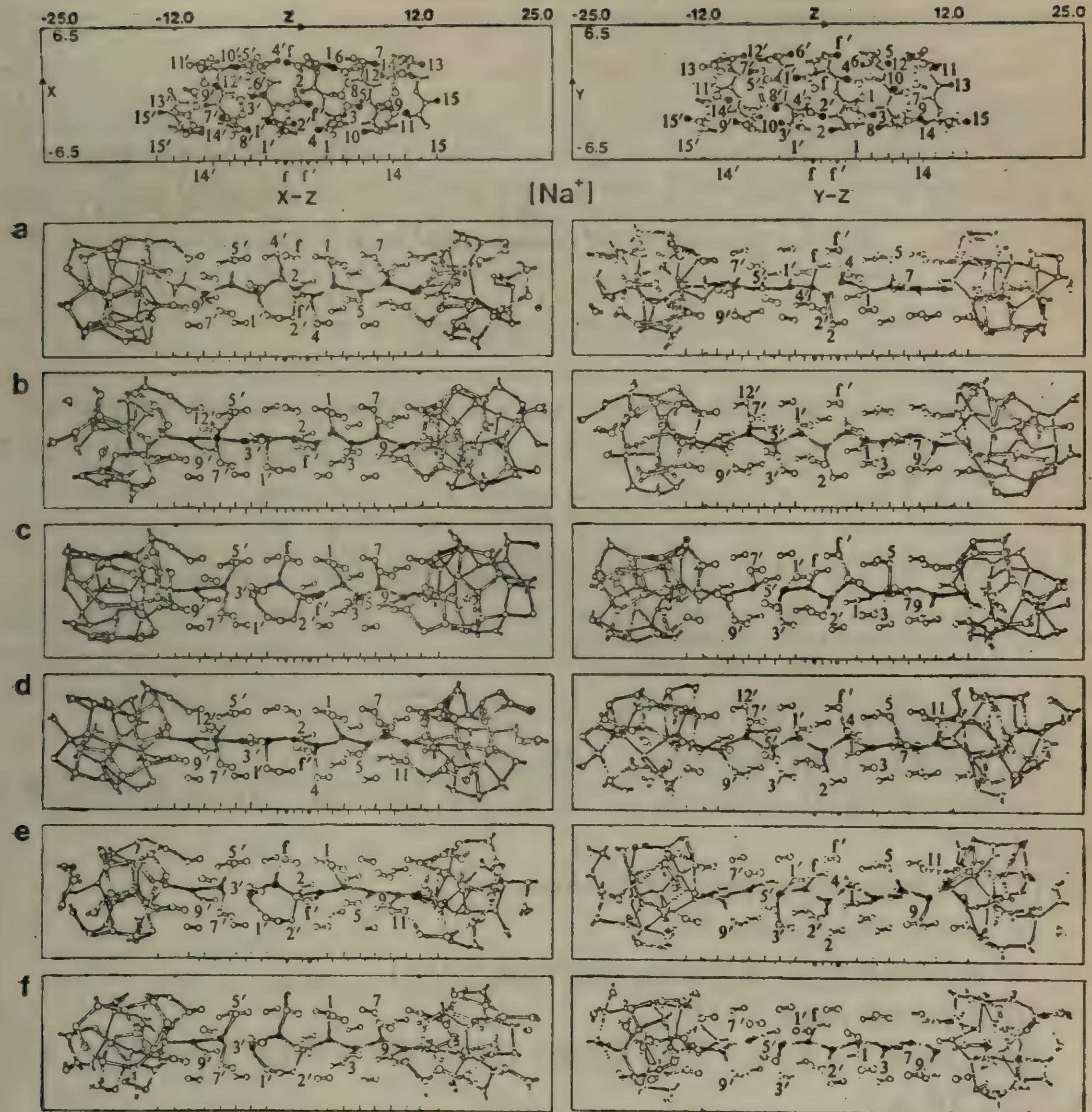


图 5 13 Na^+存在时，溶剂化短杆菌肽 A 通道内的键合结构图。顶部两张只画出了短杆菌肽 A 的骨架基团，其余的图表示 Na^+ 的平均位置处于 $a: z=-0.01$, $b: z=0.35$, $c: z=0.60$, $d: z=0.87$, $e: z=1.18$, $f: z=1.54$nm

动到通道口时，它们的溶剂化水分子个数便增加。K^+ 离子在 $z=0.79$ nm 时，阳离子的两个水化层内均有 3 个水分子，当 K^+ 运动至 1.14 nm 时，第一水化层仍含 3 个水分子，第二水化层变为 7 个水分子了。

从计算中可进一步分析溶剂化和水化的能量状况。表 5-4 总结了这方面的数据。它给出了：

表 5-4 数值分析结果

A. K^+

K^+	W1	W2	WD1	WD2	S 1	S 2	S 12	I/W	I/GA	I/T	Tot	Tot/#	STD
a	2	2	2.48	3.77	−130	−77	−207	−302	−210	−361	−6624	−81.8	.23
b	2	2	2.52	4.45	−134	−62	−196	−278	−209	−347	−6648	−82.1	.28
c	2	2	2.48	4.76	−138	−48	−186	−288	−210	−354	−6651	−82.1	.27
d	2	3	2.50	5.06	−136	−52	−190	−292	−197	−342	−6631	−81.9	.29
e	3	7	2.62	*	−184	−84	−268	−396	−169	−367	−6658	−82.2	.28
f	5	*	2.72	*	−278	*	*	−522	−187	−448	−6742	−83.2	.35

B. Na^+

Na^+	W1	W2	WD1	WD2	S 1	S 2	S 12	I/W	I/GA	I/T	Tot	Tot/#	STD
a	2	2	2.07	3.81	−187	−76	−263	−414	−278	−485	−6835	−84.4	.38
b	2	2	2.10	4.17	−188	−66	−254	−400	−294	−494	−6854	−84.6	.27
c	2	2	2.10	4.17	−188	−42	−230	−402	−287	−488	−6863	−84.7	.35
d	2	3	2.11	4.66	−188	−84	−262	−460	−270	−500	−6855	−84.6	.30
e	4	7	2.37	*	−318	−90	−408	−710	−217	−522	−6886	−85.0	.30
f	6	*	2.34	*	−443	*	*	−776	−165	−553	−6953	−85.8	.43

* 未确定

(1) 溶剂化离子第一(W 1)和第二(W 2)配位层中水分子的数目；

(2) 离子到第一和第二配位层的水分子的平均距离 WD 1、WD 2；

(3) 第一、第二配位层溶剂化能量 S1、S2，S1 和 S2 的加和；S12，总溶剂化能 I/W；单位 kJ/mol。

(4) 离子与短杆菌肽 A 的相互作用能 I/GA；

(5) 总的离子相互作用能 I/T；

(6) 系统的总能量 Tot，每摩尔水时的系统总能 Tot/#；

(7) 离子配位的标准偏差 STD。

系统的总能量 Tot 可分解为阳离子贡献部分 I/T 和水分子贡献部分 W/T，即 Tot＝I/T＋W/T；I/T 可表为

$$I/T=\frac{1}{2}I/W+I/GA,$$

W/T 可表为

$$W/T=W/W+\frac{1}{2}W/I+W/GA$$

其中 W/I＝I/W。

如上简述，无论是 Na^+ 还是 K^+，在 $z=-1.2$ nm 和 $z=1.2$nm 之间，短杆菌肽 A 的通道内有 9 个水分子，这些水分子处于线性链状构象。短杆菌肽 A 的膜通道是非常窄的，在某一瞬间，对某一截面来讲，只能通过一个阳离子或水分子。一个阳离子要穿过通道，必须推开通道内所有的

水分子。当$1.1\ nm > |z| > 0.3\ nm$时，从能量允许的区域中来看，横截面的半径对H_2O和K^+约为0.7 a.u.,对Na^+,约为1.0 a.u.。位能图中,能量最低点位于螺旋短杆菌肽A的骨架原子之间。阳离子和水分子是沿着螺旋途径在短杆菌肽A骨架原子的两链之间前进运动的。其轴半径对K^+为0.05 ± 0.01 nm，对Na^+为0.08 ± 0.02 nm。这跟K^+离子穿过通道比Na^+快的结果是一致的，因为由于Na^+运动路程比K^+要长1.4倍，并且Na^+更易停留在羰基氧附近。由此可以推断，Li^+的运动途径将更长，而Cs^+很可能是直线运动。因此Cs^+比K^+能更快通过通道，Li^+比Na^+更慢。如此很好地解释了短杆菌肽A跨膜的离子选择性大小的问题。

5.3 肽链构型的能量分析

从结构上看，肽链是蛋白质的骨架基础，多肽的构象研究对理解蛋白质的立体结构是非常重要的。下面我们介绍肽结构的描述和对规则多肽链进行研究的一些结果。

5.3.1 肽结构的描述

从化学家的观点来看，蛋白质是由氨基酸残基以—NH—CHR—CO—形式线性连接形成的。对理解多肽链的立体构象来讲，以氨基酸残基为单位的描述往往不够好。这是因为第一个氨基酸的CO与第二个氨基酸的NH之间的相互作用所形成的C—H键，即肽键，具有双键的特征。图5-14画出了导致产生C—N键双键特征的二个共振结构式。由于共振的结果，从一个α碳原子出发至另一个α碳原子，整个(C^α—CO—NH—C^α)原子集合，全部落在一个平面上。该集合具有相当的牢固性，它是研究蛋白质结构的最重要的基本结构单元，称为肽单位。

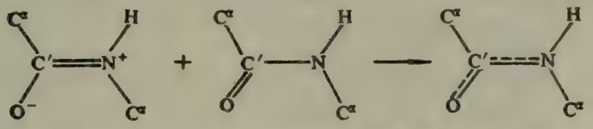


图 5-14 肽单位的两个共振结构式及其所产生的具有部分双键特征的肽键C′—N。

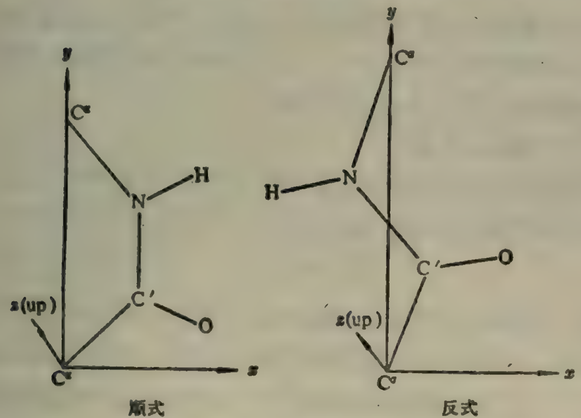


图 5-15 肽单位的顺式和反式模型

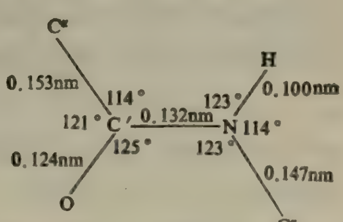


图 5-16 肽单位的标准结构数据

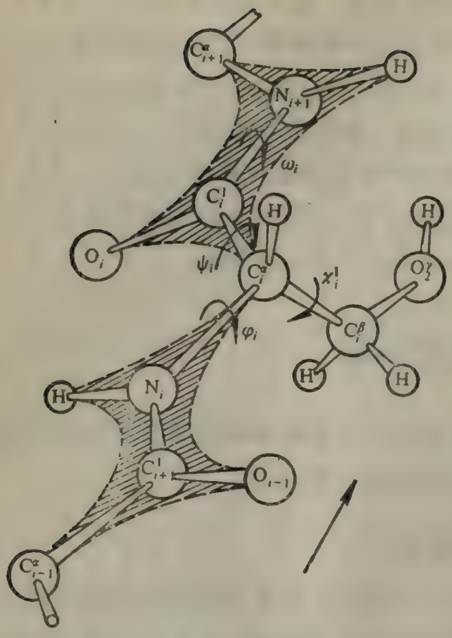


图 5-17　多肽链的构象

肽单位所在平面，又称酰胺平面。晶体结构分析表明，蛋白质中的绝大多数酰胺平面都为平面结构，只有少数情形酰胺面受到扭折。

肽单位有顺式和反式两种（图 5-15）。除个别例外，蛋白质中的肽单位全是反式的。反式肽单位通常采用的结构数据是 Corey 和 Pauling 提出的，如图 5-15 所示。

肽链可以看作是由许多肽单位在 α 碳原子上相互连接而成的。在 α 碳处连接两个肽的两个键 C^α—N 和 C^α—C 是可以自由旋转的，从这两个键的键长数据可以看出，它们都是单键。这样，抓住两个相邻的酰胺平面，以共用的 C^α 为定点旋转，如图 5-17，两个酰胺平面间可以有无数个位置。这两个肽的位置可用一对扭角，又称二面角来描述，记为(ϕ,ψ)。

假如一条肽链上所有的酰胺平面都作这种旋转，肽链就会象一条柔软的链条，构成各种不同的结构。此时，整个肽链的位置，就可用一组二面角的集合$\{(\phi_i,\psi_i)\}$来描述。也就是说，多肽链的特征构象是肽链基团绕共价单键 C^α—N，C^α—C 旋转而形成的，它们又称为多肽链的折叠。给出了一个肽段的$\{(\phi_i,\psi_i)\}$，就完全定量地给出了这个肽段的构象。对第 i 个氨基酸残基，它的 ϕ、ψ 各自包含的原子及顺序如下：

$$\phi_i(C_{i-1},N_i,C_i^\alpha,C_i)$$
$$\psi_i(N_i,C_i^\alpha,C_i,N_{i+1})$$
$$\omega_i(C_i^\alpha,C_i,N_{i+1},C_{i+1}^\alpha)$$

其中第二、第三两个原子间的共价键就是旋转的中心键。

理想的肽单位，如前所述，是平面结构。但实际上，并非所有原子全在一个平面上。这种非平面性用二面角ω来表示。它是C_i^α，C_i，N_{i+1}构成的平面与C_i，N_{i+1}，C_{i+1}^α构成的平面之间的夹角，C_i—N_{i+1}是旋转的中心键。由于肽键具有双键特征，所以 ω 值接近 180°。

ϕ_i,ψ_i,ω_i 全为零，称为全伸展构象，对应下列结构状态：

$\phi_i=0$，N_i—C_{i-1}与 C_i^α—C 成顺式或 C_i^α—C 与 N_i—C_i成反式；

$\psi_i=0$，C_i^α—N_i 与 C_i—N_{i+1}成顺式或 C_i^α—N_i 与 C_i—N_{i+1}成反式；

$\omega_i=0$，C_i^α—C_i与N_{i+1}—C_{i+1}^α成顺式。

肽单元相互旋转，使主链出现各种构象。主链构象不同，侧链 R 就出现在不同的空间位置上。侧链有大有小，相互间或者吸引，或者排斥。蛋白质分子因侧链的这种作用，使得主链与侧链相互影响、相互制约，达到某一最稳定状态。为了适应侧链构象的要求，有时要破坏主链构象的安排。然而，侧链构象的形成又将巩固主链的构象。因此，一个蛋白质分子的比较完全的描

述.除了主链的$\{(\phi_i,\psi_i,\omega_i)\}$外,还必须对侧链的构象有一个定量的描述。

侧链中原子的构象可用一系列二面角$\chi^1,\chi^2,\cdots\cdots$来描述。$\chi^1$表示绕键1,即$C^\alpha—C^\beta$键的旋转,$\chi^2$表示绕键2,即$C^\beta—C^\gamma$键的旋转,如图5-17。对Val这类与$C^\beta$有$C^{\gamma1}$、$C^{\gamma2}$连接的,则可另加数符标记,写成$\chi^{jk}$。

二面角又叫扭角或构象角。利用它,可将包括主、侧链在内的整个蛋白质分子的构象定量地表达出来。当然,对蛋白质分子主链和侧链的完备的精细的描述,仅有主链的一套二面角$\{(\phi_i,\psi_i,\omega_i)\}$和侧链的一套二面角$\{\chi^{jk}\}$是不够的,必须还有其它的一些参数,如各键键长;肽单位$C_i^\alpha-C_i-N_{i+1}-C_{i+1}^\alpha$中,与$C_i$键合的氧原子偏离酰胺平面的程度(往往用$\nu^0$表示,即$\nu^0$表示$C_{i+1}^\alpha—N_{i+1}—C_i—O_i$的非平面性);肽单位中与$N_{i+1}$键合的氢偏离酰胺平面的程度(它用$\nu^H$表示,即$\nu^H$反映了$C_i^\alpha—C_i—N_{i+1}—H$的非平面性),等等。但毫无疑问,描述蛋白质空间结构最主要的参数是肽链的一组二面角$\{(\phi_i,\psi_i,\omega_i)\}$和侧链的二面角$\{\chi^{jk}\}$。

为了描述蛋白质多肽链的性质,Ramachandran以构象角ϕ为横坐标、构象角ψ为纵坐标,作出一直方图,其中的一点对应一对(ϕ,ψ)。将一个蛋白质所有肽段的(ϕ,ψ)画在图上,一条肽链的构象就鲜明地显示在一张图上,这种(ϕ,ψ)图就称为Ramachandran图(195页图5-18),简称拉氏图。从拉氏图中,能得到蛋白质多肽链的折叠特征(图5-17)。

表 5-5 蛋白质中非键原子的允许距离(nm)

第一行是一般可接受的范氏距离,括号中是拉氏在小分子结构中发现的最小允许距离

	C	N	O	H
C	0.320 (0.300)	0.290 (0.280)	0.280 (0.270)	0.240 (0.220)
N		0.270 (0.260)	0.270 (0.260)	0.240 (0.220)
O			0.270 (0.260)	0.240 (0.220)
H				0.200 (0.190)

5.3.2 规则多肽链分子构型的能量分析

肽链中构型能量的计算一般取Momany等提出的势函数和参数值。两个单体间的相互作用能包括静电作用、色散吸引和非键排斥。氢原子(羟基和胱氨上的氢和氧原子以及氨基氮原子之间的非键作用采用广义的氢键函数。

$$\Delta E_{AB}=\sum_{i\in A}\sum_{j\in B}\{q_iq_j/\varepsilon_{ij}r_{ij}+[A_{ij}/r_{ij}^{12}-C_{ij}/r_{ij}^6]$$
$$+[A'_{ij}/r_{ij}^{12}-B'_{ij}/r_{ij}^{10}]\}$$

式中A_{ij}、C_{ij}为非氢键原子之间的势参数,A'_{ij}、B'_{ij}为氢键原子之间的势参数。

表 5-6　蛋白质二级结构及一些规则结构的特征(ϕ,ψ)值

结　　构		ϕ(度)	ψ(度)
右手 α 螺旋		−57	−47
3_{10} 螺旋		−49	−26
平行 β 折叠层		−119	+113
逆平行 β 折叠层		−139	+135
回　折	I 型	$\begin{pmatrix}-60\\-90\end{pmatrix}$	$\begin{pmatrix}-30\\0\end{pmatrix}$
	II 型	$\begin{pmatrix}-60\\80\end{pmatrix}$	$\begin{pmatrix}120\\0\end{pmatrix}$
	III 型	$\begin{pmatrix}-60\\-60\end{pmatrix}$	$\begin{pmatrix}-30\\-30\end{pmatrix}$
胶原三股螺旋		−51	+135
		−76	+127
		−45	+180
聚甘氨酸 II		−80	+150
聚 L-脯氨酸			
	I	−83	+158
	II	−78	+149

总的构型能包括不同单体间的相互作用能、侧链与骨架间的相互作用能和转角势能：

$$E_{\text{Tol.}}=\sum_{i=1}^{m}\Delta E_i+0.8/2\{2-\cos 3\phi+3\cos\psi\}+E_{s\leftrightarrow m}$$

侧链跟骨架的作用能 $E_{s\leftrightarrow m}$ 按相同的方法计算。介电常数在整个计算中取为 4。

势能参数 A、C 由下式给出：

$$A=1/2\cdot C\cdot R_e^6$$

$$C=3/2\cdot\frac{eh}{\Delta\pi m_e}\cdot\alpha_A\alpha_B/[(\alpha_A/N_A)^{1/2}+(\alpha_B/N_B)^{1/2}]$$

其中A、B表示两个原子，e、h和m_e为基本常数，N为有效价电子数。

势能参数列表 5-7 和表 5-8 中。

表 5-7　肽链构型能量参数

参数＼原子	肽氨H	羟基H	碳H	羟基O	羟基O	肽氨N	饱和C	羟基C
α	0.42	0.42	0.42	0.84	0.59	0.93	0.93	0.93
N_e	0 85	0 85	0.85	7.00	7.00	6.10	5.20	5.20
r_e	2.68	2.83	2.92	3.12	3.12	3.24	4.12	3.74
re′	2.40	2.40	2.40	3.04	3.04	3.10	3.40	3.40
q	0.27	0.22	0.0	−0.417	−0.32	−0.30	0.0	0.45

表 5-8　氢键函数能量参数　　（单位：kcal/mol）

受　体	羟基氢		羟基氧		氨　氮	
氢	A′	B′	A′	B′	A′	B′
胱氨氢	12040	4014	13033	4000	32987	8244
羟基氢	13044	5783	13033	4610	32987	8300

图 5-20 绘制了在整个 ϕ 和 ψ 空间中肽链螺旋参数 θ（以实线表示）和步高（以虚线表示）的变化曲线。这些曲线表示了肽链螺旋结构的如下特点：

1. 构型空间具有中心对称性。这是肽平面几何构型的反映；

2. 肽链所能形成的螺旋构型比较有限，它不能生成每圈螺旋中单体数目大于 5 的松散缠绕螺旋，只能形成 n 比较小的紧凑螺旋；

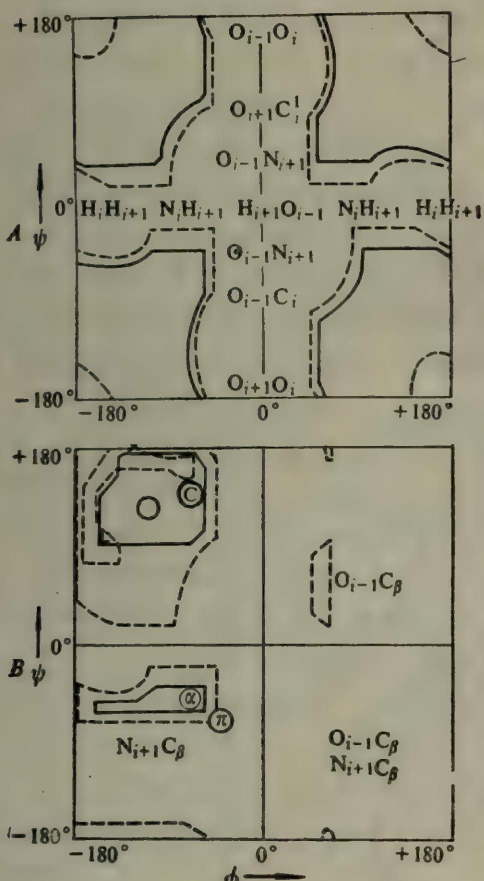


图 5-18　Ramachandran 构象图

实线范围是一般允许区，虚线范围是最大允许区。A. 是对甘氨酸的结果，B. 是有 C^{β} 的结果，有空间障碍的原子也标在图上的相应位置。α_L-左手 α 螺旋；β-β 折叠层；π-π 螺旋；c—胶原三股螺旋。

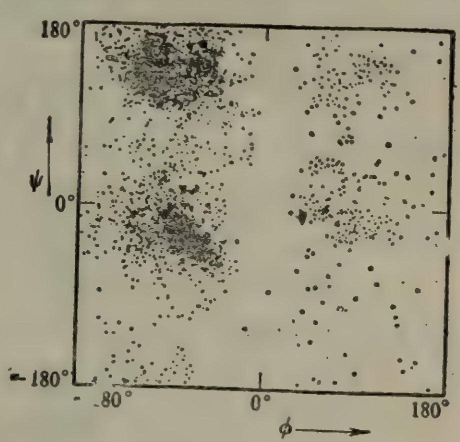


图 5-19　13 个蛋白质 2500 个残基构象在拉氏图上的分布

3. 螺旋构型区域和伸展构型区域分布得比较规则，它们几乎分布在相互平行的区间内。

图 5-21 是根据上述表中所示势参数算得的聚甘氨酸构型能量图。

粗略观察给人的印象是图 5-20 与图 5-21 有一定的对应关系。螺旋参数曲线以及它们的适

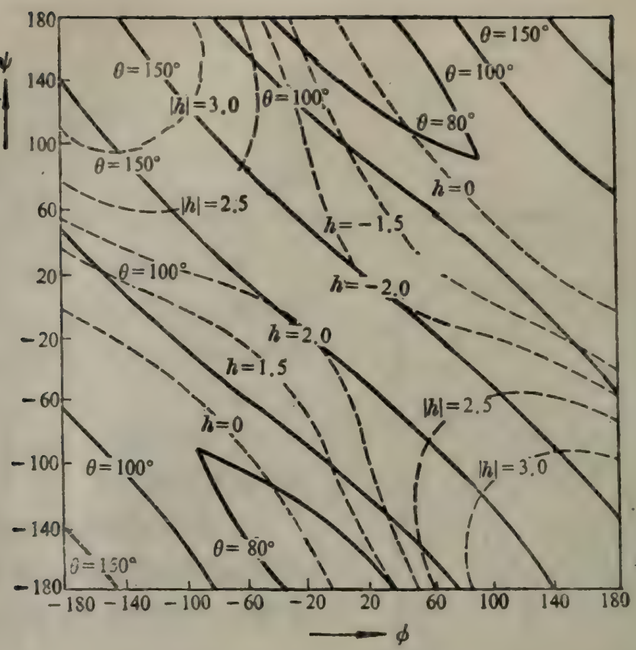


图 5-20 肽链构型螺旋参数曲线

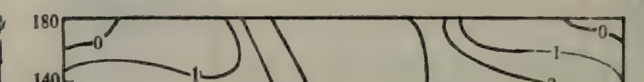


图 5-21 聚甘氨酸的构型能量图

当交叉在一定程度上勾划了构型能量曲线最外部的轮廓线。其中最明显的是下面四个区域：

1. 左上角和右下角，$h=2.0$ 和 $h=-2.0$ 两条曲线近似地描述出折叠构型低能量区的轮廓；

2. $\theta=100^\circ$ 和 $\theta=150^\circ$ 两条曲线近似地描绘出螺旋构型区最外部的轮廓；

3. 左下角 $h=0$ 的曲线限制了伸展折叠区域的轮廓；

4. $\theta=80^\circ$ 的两个半闭合曲线恰好是一个高能量区的轮廓。

这种整齐和有规则的构型在其它类型的聚分子构型中是少见的。

聚甘氨酸构型能量图中有三对对称的能量极小点。其中最深的能量极小值点的 ϕ、ψ 与标准 α-螺旋结构的 $\phi=\pm48^\circ$，$\psi=\pm57^\circ$ 基本一致。另一能量极小值点对应于肽链的伸展构型，两个相邻肽单位相对扭曲 180°，但不完全伸展。决定这个能量极小值点的主要贡献是相邻单位的相互作用。这个构型是二肽最稳定的构型。第三个能量极小点是肽链形成最宽股螺旋的构型，能量较低的原因是单体间的色散吸引能造成。

β-折叠没有出现在构型能量极小值区域，这是因为使 β-折叠结构稳定的原因是肽链间的相互作用。显然，孤立肽链构型不包括这一部分能量贡献。从单链构型能量考虑，β-折叠的能量比 α-螺旋构型能量要高 6 kcal/mol 残基。如果它们的构型熵基本相同的话，那么只有在链间相互作用能约为 6 kcal/mol 残基时，β-折叠结构才有可能与 α-螺旋构型相竞争。

图 5-22 是根据构型能量图 5-21 中静电作用和氢键作用绘制的肽链构型空间中氢键生成的区域。图 5-22 与图 5-21 最值得注意的差别是它们的能量极小值位置不同。最强的氢键所对应的构型区为 3_{10} 螺旋构型。氢键是在 1→3 两个肽单位间生成。但在总的构型能量图中，没有这个能量极小值点，原因是在这个构型中，相邻单体处于非键排斥构型。

α-螺旋构型是第二个氢键生成区域。氢键是在 1→4 单体间生成。由于在 α-螺旋构型中原子间非键作用也比较有利，所以在总构型能量中，α-螺旋也是最低能量的构型。

第三个氢键生成区是 2_7 带结构。氢键是在 1→2 两个单体间生成。这个构型也是总构型能量图中一个能量极小值点，但总构型能量图中第三个能量极小值点没有氢键生成。

为了区分肽链中氢键作用区域与静电作用的差别，图 5-23 绘出了肽链中的静电能构型图。

静电作用能主要是肽单体中的偶极子相互作用而产生的，它的分布比较平缓，一般形成 ±1 kcal/mol 残基的能量差别。可以注意到的是 α-螺旋构型和 3_{10} 螺旋构型都是静电作用比较有利的构型。

当不考虑侧链时，肽单位具有平面对称性，因而肽链骨架的构型能量图具有中心对称性。若有了侧链，肽单体失去了对称性，分为 D-和 L-两种异构体。自然界中的氨基酸多为 L-异构体。侧链对骨架构型能量的影响相当大，当 L 位的氢原子被甲基取代时，不仅低能区大大减小，而且左手和右手螺旋区形成了较深的相对能量极小值点。图 5-24 是聚 L-丙氨酸构型能量图（侧链二面角固定）。

从图中可以看到，左手螺旋能量极小点与标准 α-螺旋构型基本一致，右手螺旋能量极小点稍稍偏离标准 α-螺旋。左手螺旋和右手螺旋中侧链与骨架的相对位置不同，造成了侧链与骨架的相互作用能不同，但是对于仅仅手性相反的两种螺旋，其侧链间的相互作用必然相同。从图中

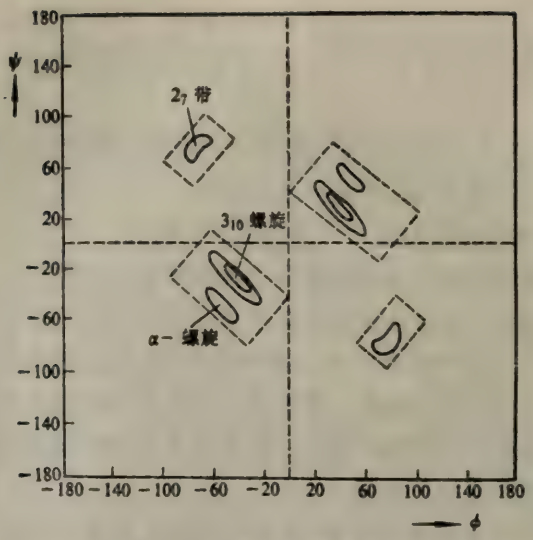


图 5-22 肽链构型中氢键区域

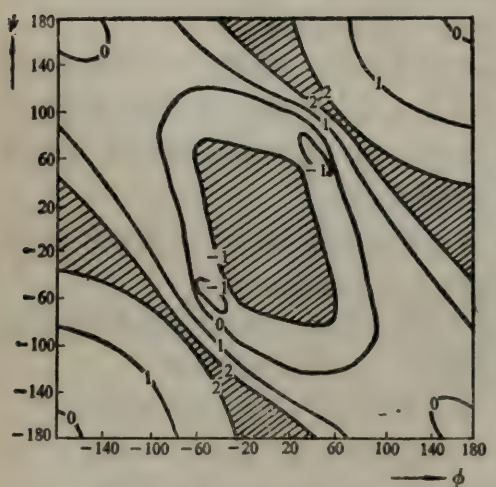


图 5-23 肽链构型中静电作用能

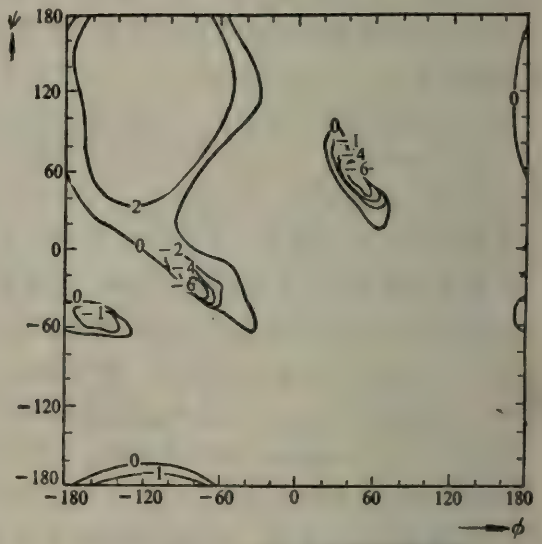


图 5-24 聚L-丙氨酸构型能量图

可以看到，骨架与侧链的非键作用对右手螺旋更为有利些，它的能量要低 2 kcal/mol 残基左右。

考察聚 L-门冬氨酸的构型能量图可讨论侧链与骨架的静电作用，这里就不再详细讨论了。

综上所述，可得多肽规则螺旋结构的如下特点：

1. 只能形成较为紧凑的螺旋，最稳定的螺旋构型是左手和右手 α-螺旋；

2. 氢键和静电作用的能量极小值点并不与非键作用的能量极小点完全重合，这种偏离对螺旋结构的稳定性是不利的；

3. 对于所有的聚-L-氨基酸，侧链与骨架的非键作用更有利于右手 α-螺旋；

4. 侧链与骨架的静电作用可能有利于左手α-螺旋,也可能有利于右手α-螺旋。

5.3.3 聚L-丙氨酸的ab initio计算

怎样表示生物大分子和底物分子周围空间各处的分子静电势是很重要的。如果某种表达方式有利的话,就便于我们认识生物分子的结构和反应活性。对于核酸碱基和碱基对,这种静电势可表达在这种平板状分子的平面上以及另一些与之平行的平面上。在螺旋状的B-DNA分子中,可以用平行于碱基对的平面和垂直于螺旋轴的平面来表示。然而,对于任何缺乏对称性的生物大分子,可用计算其"外壳"上的静电势,即分子表面静电势,从静电场角度来了解一个反应物分子和底物分子是如何接近的。图5-25 *A*表示了如何定义"外壳"并计算其上的静电势。

我们计算了有5、10、15、20、25个丙氨酸基的聚L-丙氨酸多肽α-螺旋的N端和C端的静电势。结果表明,随着螺旋长度的增加,围绕N端的静电势变得更正,而C端的静电势变得更负。此外,还可见不同螺旋长度时N端表面静电势的最大值和C端表面静电势的最小值。应该注意到,这两个极值在量值上随着螺旋长度增加得非常类似。这和有人认为的,α-螺旋可以看成是一个单个偶极,它由许多小偶极组成,每个小偶极是一个肽基,并几乎正好和螺旋轴在同一条线上,是相一致的。如图5-25所示。图5-25 *A*中是15个基组成的中等长度螺旋的静电势。六种用交叉格画成的阴影格表示了六种从最正值排列到最负值的静电势大小,旁边并附有其kcal/mol单位的值。在这种表面上,静电势的分布大致上是辐射形的,在量值上是随着距离从一个最大值点逐渐减小。对于$(\alpha|\alpha)_{15}\alpha$-螺旋这个最大值是53.5kcal/mol,最小值是−12.2kcal/mol,分别用*X*和*M*表示,可以看到,图5-25 *A*表面上可以分成44个小区域,每一个这样的区域表示大分子中贡献于这个表面外壳的一个原子球表面的可见位置。因此,在外壳上任何点上的静电势可以和这点相应的某个原子相联系,这个原子最接近这点。例如,用这种方式我们就可以知道,N端表面的最大值在最接近螺旋端部的第三个氢原子即质子上。这个氢原子在图5-25 *B*中用箭头指示出。图5-25 *B*是从N端看到的α-螺旋分子图(为了更清楚地看到α-螺旋的形式,在这里仅包括了头五个丙氨酸基)。那些用小圈及字母H、C、N或O表示的原子是那些从窗口所见贡献于这个表面分子外壳的原子球,其围绕的界限相应于图5-25 *A*中所见的N端的表面外壳。图5-25 *C*是N端表面外壳的一种三维表示,箭头标出了图5-25 *A*中看到的最大值位置。可以看出,图5-25 *C*中的表面是由二维图畸变而形成,这套"网眼"大小相应于用于计算表面外壳静电势时点的密度,它也解释了外壳边界线或原子边界线不是平滑的。表面外壳形状的分辨率受到网眼的精细限制。

图5-26 *A*表示了$(\alpha|\alpha)_{15}$聚丙氨酸多肽α-螺旋C端表面的静电势。如前所讨论,图5-26 *B*和图5-26 *C*分别表示了C端的分子图和其外壳表面的三维表示。C端表面静电势的分布和N端类似,大致上是辐射形的。但这时的静电势从居中的最小值随着距离越来越正,这和N端的中心最大值正好相反。C端表面的最小值为−61.3 kcal/mol,而最大值则为21.0 kcal/mol。图5-26 *A*中亦分别用*M*和*X*标出。C端静电势最小(负)值位于分子端部起第三个羰基的氧原子的表面。图5-26 *B*中用箭头指出了这个原子,图5-26 *C*中则标出了三维表面上的这个最小值位置。

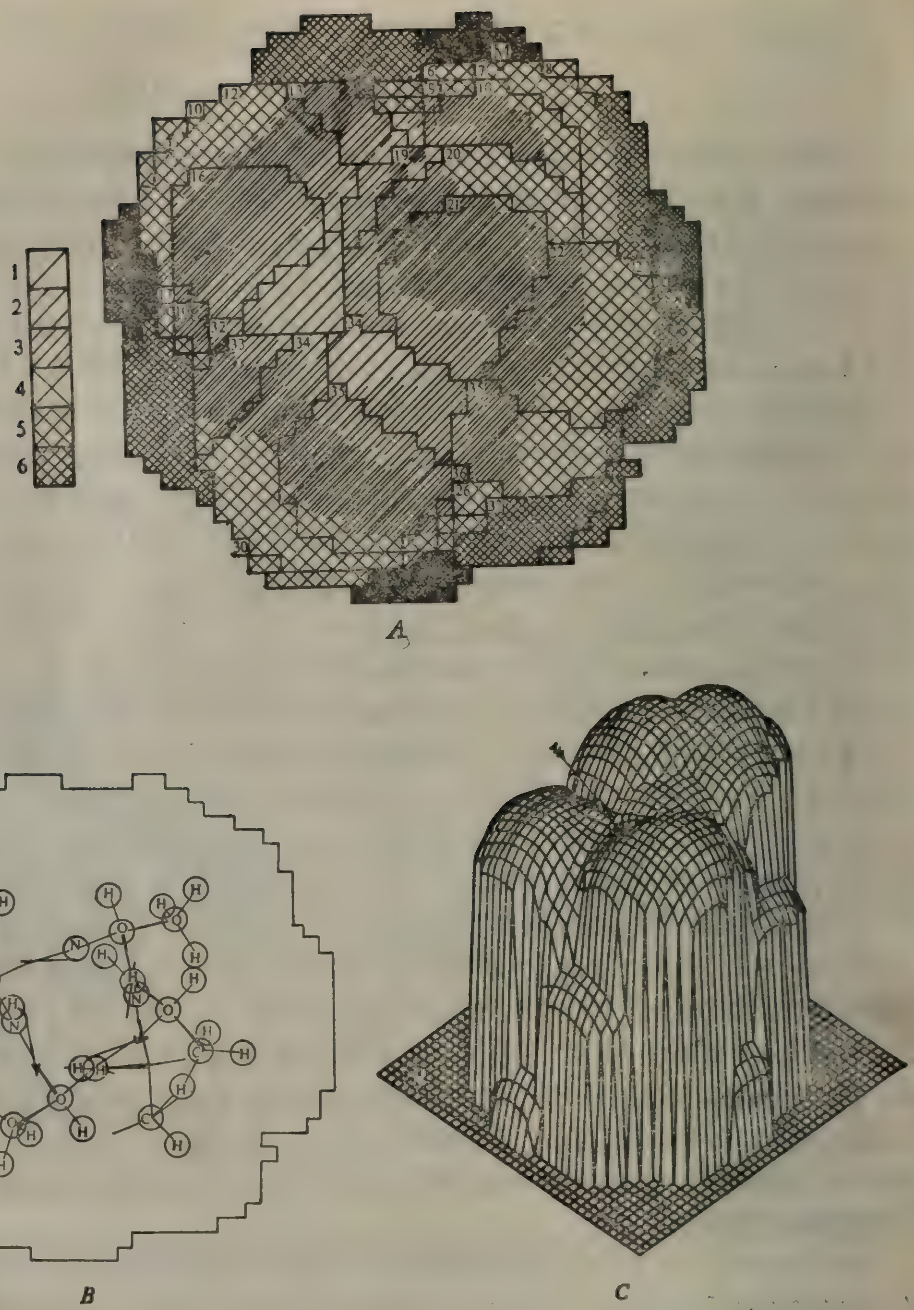


图 5-25　*A*. 聚 L-丙氨酸多肽 *α*-螺旋N端表面静电势，1. 54.00 到 43.00，2. 43.00 到 34.00，3. 34.00 到 24.00，4. 24.00 到 15.00，5. 15.00 到 5.00，6. 5.00 到 −5.00。
B. 聚 L-丙氨酸多肽 *α*-螺旋N端表面分子结构图示。
C. 聚 L-丙氨酸多肽 *α*-螺旋N端表面外壳的三维图示[18]。

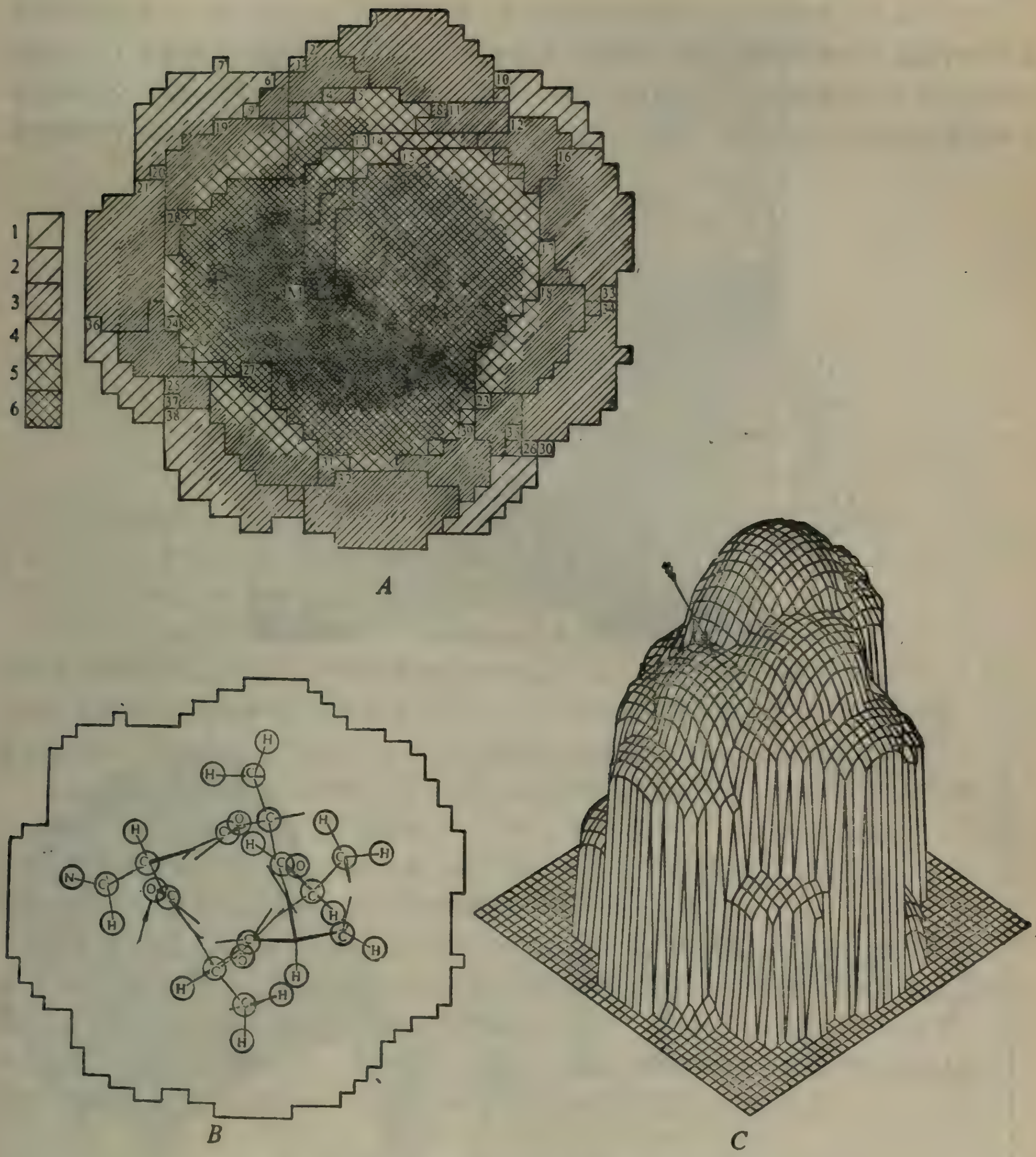


图 5-26 *A* 聚-L-丙氨酸多肽 α-螺旋C端表面静电势

1.11.00 到 −1.00 2. −1.00 到 −13.00 3. −13.00 到 −25.00

4. −25.00 到 −37.00 5. −37.00 到 −49.00 6. −49.00 到 −62.00。

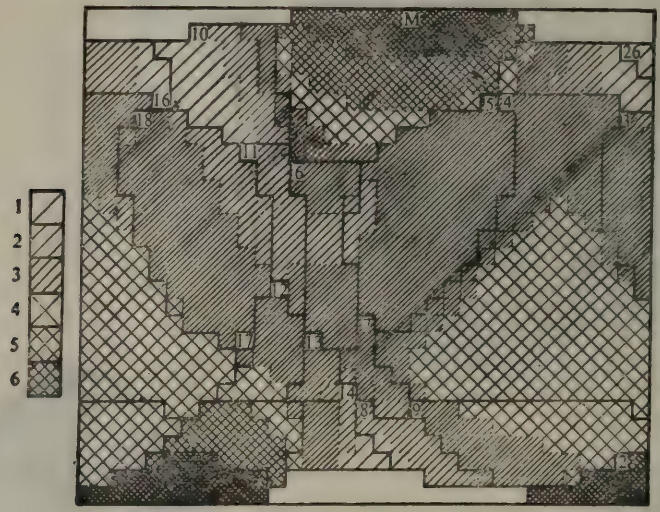
B 聚-L-丙氨酸多肽 α-螺旋C端表面分子结构图示

C 聚-L-丙氨酸多肽 α-螺旋C端表面外壳的三维图示

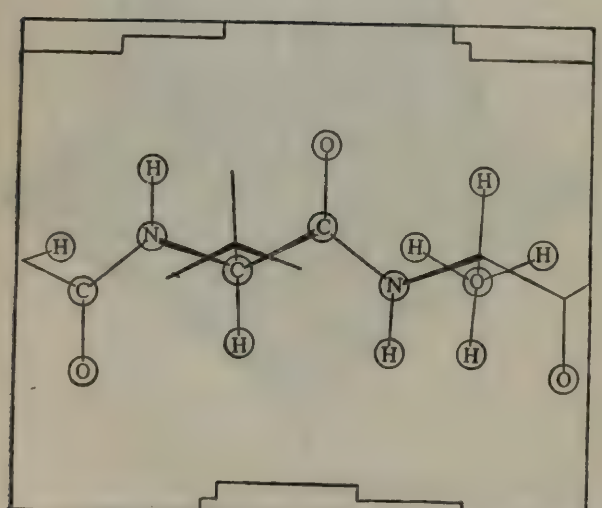
所有研究过的包含有从 5 个到 25 个丙氨酸基的 α -螺旋在N端和C端都有非常类似的表面静电势分布。所不同的仅仅是其两端表面最大值和最小值的量值。因此，本文仅列出中等长度

螺旋的图。

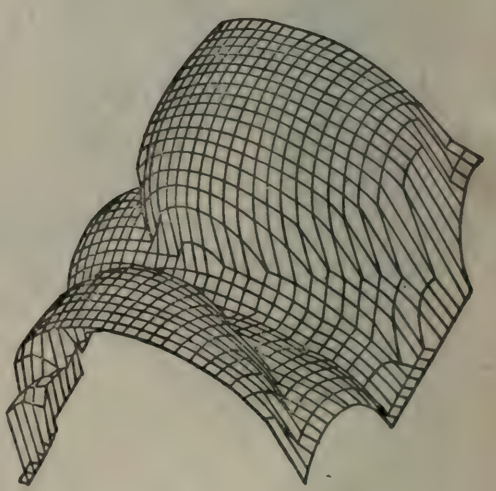
为了计算在反平行β-片层中每个丙氨酸基上的静电势分布，我们先研究了具有这种构象的单个聚丙氨酸多肽链但长度不同时的中心丙氨酸基上的表面静电势。我们在此选取了一个足够大的包括两个中心残基的窗口，因这样就可看到残基的两边。在β-片层中链上下一个丙氨酸是反向重复排列的，仅仅是绕轴反转 180°，我们分别计算了长度为 4、6、8、10、14 和 22 个丙氨酸基



A



B



C

图 5-27 *A* 聚-L-丙氨酸多肽反平行 β-片层单链时中心丙氨酸基的表面静电势
1、6.00 到3.00 2、3.00 到 1.00 3、1.00 到 0.0 4、0.0 到 −1.00
5、−1.00 到 −3.00 6、−3.00到−6.00
B 聚-L-丙氨酸多肽反平行 β-片层单链时中心丙氨酸基的分子结构图示
C 聚-L-丙氨酸多肽反平行 β-片层单链时中心丙氨酸基表面外壳的三维图示



图 5-28 聚-L-丙氨酸多肽反平行 β-片层中心链中心丙氨酸基的表面静电势。
1.16.00 到10.00 · 2.10.00 到 6.00 3.6.00 到 1.00
4.1.00 到 −2.00, 5. −2.00 到 −7.00 6. −7.00 到 −12.00

的链。计算结果表明，链长超过 10 个丙氨酸基时其表面静电势就几乎没有什么变化了。图 5-27 *A* 是一个有 10 个丙氨酸基的链的中心部位的静电势分布。如前一样，图 5-27 *B* 和图 5-27 *C* 则分别是这个部位的分子图和其分子外壳表面的三维表示。

图 5-27 *A* 和图 5-27 *B* 的比较表明，表面静电势极小值（−11.9 kcal/mol）位于中心氧原子表面，而极大值(15.4 kcal/mol)则位于邻近肽键的 NH 的质子上。这个表面的其余部分主要由 C、CO 中的 C′ 和丙氨酸的甲基等原子所占据。朝上方向的甲基在图 5-27 *C* 中非常明显，静电势在这些区域相当弱，大致变化在 −5—10 kal/mol 的范围内。

我们现在讨论真正的聚丙氨酸多肽反平行 β-片层。假设它由 10 个丙氨酸基的链组成，因为从前面讨论中可见它也能代表更长的链了。我们计算了由 3 、5 、7 和 9 条链组成的 β-片层的中心链中心丙氨酸基上的表面静电势。这个计算中所用的窗口加宽了一些，以使我们不仅能看到中心链而且还能看到邻近的链，因而更能看清楚 β-片层的结构。

图 5-28 列出了其中最大一个片层中心丙氨酸基上的表面静电势。我们发现，在片层中邻近链的效应是在量值上减小了丙氨酸基氧原子附近的负静电势，同时也减小了NH 质子上的正静电势。因此，单个$(Ala)_{10}$链中所得的最大和最小静电势 15.4kcal/mol 和 −11.9kcal/mol，在 9 个链的片层中则分别减小至 5.7 kcal/mol 和 −5.6 kcal/mol，片层中的最小值位置和在单个链中一样，位于中心羰基氧上，但最大值位置却移至中心甲基的一个质子上。

5.4 蛋白质分子的量子化学研究

蛋白质分子的分子量一般在 10^4以上，要对它们进行哪怕是最简单的量子化学或分子力学计算在目前都是不现实的。蛋白质分子的量子化学研究方法有两种，一是“分割法”，将蛋白质分子按一定的方式分割为“碎片”，把对蛋白质分子的量化计算变换为对“碎片”的计算；二是模型法，抽象所研究体系的特点，构造模型，把对蛋白质分子的计算化为对模型分子的计算。木瓜蛋白酶

和胰岛素的量子化学研究属于第一种研究方法，下一章介绍的酶催化机制的量化研究采用的则是模型法。

5.4.1 胰岛素的量子化学研究

胰岛素是第一个被阐明化学结构的蛋白质。

1955 年，Sänger 等利用酶和化学相结合的方法成功地搞清了牛胰岛素的全部化学结构。牛胰岛素是由两条肽链组成的，一条叫 A 链，另一条叫 B 链。A 链由 21 个氨基酸组成，B 链由 30 个氨基酸组成。A、B 两条肽链通过两个双硫键连接起来。另外，A 链本身还有一个链内双硫键。它的分子量为 5 734。其化学结构如图 5-29。

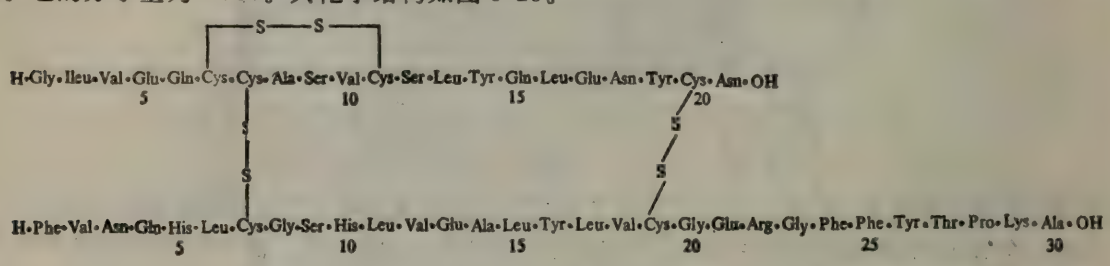


图 5-29 牛胰岛素的化学结构

此后，Sänger 以及其它科学工作者又测定了不同动物来源的胰岛素的化学结构。虽然它们都具有相同的连接方式，但其组成的氨基酸都可以有很大的不同。约有一半以上的氨基酸残基是可以改变的(表 5-9)。

1926 年，Abel 首先制得了结晶的纯胰岛素样品，为胰岛素的研究工作提供了有利条件。60 年代末—70 年代初，Hodgkin 和我国科学工作者分别独立地用 X 光衍射方法完成了高分辨率的胰岛素结构分析工作，使人们对结构复杂的胰岛素分子的各个基团甚至某些原子在空间所处的位置有了比较全面的了解。近几年来，更高分辨率的报告不断出现。各国科学家还对胰岛素的全合成、生物合成、分泌、贮存、生理功能、结构与性能、化学修饰、作用原理等方面进行了大量的研究，以胰岛素为主题的国际会议已举办过多次。在这节里，我们主要讨论如何利用计算机和量子化学方法对胰岛素分子进行研究。

5.4.1.1 胰岛素分子的疏水内核和分子的几何特征

实验表明，几乎所有的球蛋白分子都有一个疏水内核，而亲水性的基团往往暴露在分子表面。若蛋白质分子比较小，本身难于满足这一规则时，可以形成二聚体等缔合物，使缔合物的核成为疏水核。要弄清一个由成千上万个原子组成的蛋白质分子的疏水和亲水基团的分布状况，除了根据结构数据搭建分子模型外，目前尚无其它方法。

我们定义分子表面为该分子中的每一个原子球面的集合体所形成的几何曲面。显然，对构成分子表面有“贡献”的原子应处于分子的表面部分，而无“贡献”者则处于分子内部。这种定义方法是和用于计算大分子原子空间装配性的 CPK 模型相当的。

上述定义并没有指明原子半径的取法。在具体计算时，根据需要，原子的球面半径可以取

表 5-9　不同动物的胰岛素 A 链氨基酸排列顺序

种　属	①	②	3	4	⑤	⑥	⑦	8	9	10	⑪	12	13	14	15	⑯	17	18	⑲	⑳	㉑
牛	Gly	Ileu	Val	Glu	Gln	Cys	Cys	Ala	Ser	Val	Cys	Ser	Leu	Tyr	Gln	Leu	Glu	Asn	Tyr	Cys	Asn
人、猪、狗、兔、抹香鲸								Thr		Ileu											
鲸(Sei Whale)										Thr											
北京鸭								Glu	Asn	Pro											
山羊、马									Gly												
象								Thr	Gly												
大　鼠				Asp				Thr		Ileu											
豚　鼠				Asp					Gly	Thr		Thr	Arg	His				Ser			
小　鸡								His	Asn	Thr											
鱼(Toad fish I)								His	Arg	Pro		(Asp	Ileu)	Phe	Asp		Gln	Ser			
鱼(Toad fish II)								His	Arg	Pro		(Asp	Lys)	Phe	Asp		Gln	Ser			

"○" 表示不变的氨基酸　　(　) 表示未完全肯定的氨基酸

不同动物的胰岛素 B 链氨基酸排列顺序

种　属	0	1	2	3	4	⑤	⑥	⑦	⑧	9	10	⑪	⑫	13	14	⑮	⑯	17	⑱	⑲	20	21	22	㉓	㉔	㉕	㉖	27	28	29	30
牛、猪、马、山羊、狗、抹香鲸、鲸(Sei Whale)		Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr	Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Ala
大鼠 I				Lys																											Ser
大鼠 II				Lys																										Met	Ser
豚　鼠				Ser	Arg						Asn				Thr			Ser			Gln	Asp	Asp					Ileu			Asp
小　鸡		Ala	Ala																									Ser			
人　象																															Thr
琵琶鱼	Val	Ala	Pro(Ala,Gln)											(Asp)								(Asp)						(Asn)			
鱼(Toad fish I)	Met(Ala,Pro,Pro)													Asp								Asp						Asn			
鱼(Toad fish II)	Met(Ala,Pro,Pro)													Asp								Asp						Asn	Ser		
北京鸭		Ala	Ala																									Ser			Thr

"○"表示不变的氨基酸　　(　)表示未完全肯定的氨基酸

原子的共价半径(RC)、Van der Waals 半径(RV)或 Van der Waals 半径的若干倍(RV-K)。可以预计，球面半径取值不同，所得分子的表面和内核情况也不会相同。这是因为不同的半径反映了不同种类的相互作用范围。也就是说，对不同的相互作用来讲，分子的表面和内核是应有所区别的。

根据作者自行设计的生物大分子表面和内核分析程序，分别以 RC、RV、RV-1.7 为原子的球面半径，对三方二锌（猪）胰岛素晶胞不对称单位中的两个单体：单体分子 I 和单体分子 II 及其二聚体进行了计算，并分析了内核和表面的组成情况。

计算中所用胰岛素分子的坐标数据取自 Hodgkin 分辨率为 0.19nm 的晶体结构测定。

表 5-10 是胰岛素两个单体分子和二聚体在各种半径值下分子内核的组成情况。

表 5-10　胰岛素分子的内核组成

RC	分子 I	
RC	分子 II	
RC	二聚体	
RV	分子 I	A_6(CA、C、H 4)A_{16}(CD 2、H9、H 10、H 11)B_{15}(CB、CD 2、H 6、H 7、H 8)
RV	分子 II	A'_6(SG、O)A'_{16}(H 4、H 7)B'_{15}(C、H 3、H 4、H 5)
RV	二聚体	A_6(CA、C、H 4)A_{16}(CD 2 H 9、H 10、H 11)B_{12}(CG 1、H 4、H 5)B_{15}(CD2、H 7、H 8)B_{24}(H 2、H 5、H 6)B_{26}(CY、OH、CE 2、CD 2、O) A'_6(SG、O)A'_{16}(N、H4、H 7)B'_{11}(CD 1、H 3 H 9)B'_{12}(CG 1、C、H 6、H 7、H 8) B'_{15}(C、H 3)B'_{23}(H 2)B'_{24}(O、H 2)B'_{26}(OH、H 1、H 7、H9)
RV-1.7	分子 I	A_6、A_{11}(CA、H 2)A_{16}、B_6(CD 1、H 2、H 7、H 8)B_{11}、B_{15}、B_{19}(C、O、H 2)
RV-1.7	分子 II	A'_2 A'_6 A'_7(SG、O)A'_{11}(H 2)A'_{16} A'_{20}(O、H4)B'_6(CD 2、H 2、H 10、H 11)B'_{11} B'_{15} B'_{19}(C、O)
RV-1.7	二聚体	A_6 A_{11}(CA、H 2)A_{16} B_6(CD1、H2、H 7、H 8)B_{11} B_{12} B_{15} B_{19}(CD、H2)B_{23} B_{24} T_{26}(OH、CE 2、H 7) A'_2 A'_6 A'_7(SG、O)A'_{11}(H2)A'_{16} A'_{20}(O、H 4)B'_6(CD 2、H 2、H 10、H 11)B'_{11} B'_{12} B'_{15} B'_{19}(C、O)B'_{23} T'_{24} T'_{26}(OH).

注：括号内的原子为在分子表面者。

从表中可以清楚地看到：当原子半径取不同值时，胰岛素单体及二聚体的分子内核和表面组成，正如所预料的那样，有很大的差异。

当球面半径为 RC 时，几乎所有的氢原子和绝大部分的非氢原子都参加了分子表面的构造。因此，对 RC 面来讲，不存在分子内核问题。

当球面半径为 RV 时，有若干残基，它们的许多组成原子都隐蔽了起来。这些残基是两个单体中的 A_6-胱氨酸，A_{16}-亮氨酸和 B_{15}-亮氨酸。

当球面半径为 RV-1.7 时，上述氨基酸的全部原子以及 B_{11}-亮氨酸的全体，B_{19}-胱氨酸，A'_7-胱氨酸、A'_{11}-胱氨酸、B_6-胱氨酸、B'_6-胱氨酸，B'_{19}-胱氨酸的大部分原子均不在分子表面(不带“′”的为单体分子 I，带“′”的为单体分子 II)。注意到上述残基的疏水性，可以看到，胰岛素与其它球蛋白分子一样，存在着疏水内核，其核心部分为单体的 A_6-胱氨酸、A_{16}-亮氨酸、B_{11}-亮氨酸。而 B_{15}-亮氨酸、A_{11}-胱氨酸、B_6-亮氨酸和 B_{19}-胱氨酸也部分参与。这与基于搭建分子模型而得到的结论相一致。

两个单体形成二聚体后，对 RC 面来讲，情况没有发生多大变化。

对 RV 面来讲，B_{12}-缬氨酸、B_{15}-亮氨酸、B_{23}-甘氨酸、B_{24}-苯丙氨酸、B'_{15}-亮氨酸、B'_{23}-甘氨酸、B'_{24}-苯丙氨酸、B'_{26}-酪氨酸的不少原子都不在分子表面了。

和聚合前的情况相比，我们再次注意到了这些氨基酸的疏水性。可以断言，两个单体分子是在各自 B 链的第 12、第 23、第 24 和第 26 号氨基酸所在区域发生聚合的。聚合后，这些氨基酸自然成为二聚体的疏水内核。这是球蛋白缔合产生缔合体疏水内核的一个具体例子。

蛋白质分子表面经常会出现一些"洞穴"。这些"洞穴"常常是其活性中心所在的地方。例如许多水解酶，它们的活性中心所在的地方多为"洞穴"结构。若水解酶作用的底物是大分子，这种"洞穴"便成为一种凹槽。因此了解大分子表面的几何特征，即其表面的形貌，对认识活性部位和其环境是有意义的。

根据分子表面的定义，从任何方向观察到的生物大分子的表面形貌，我们都可以方便地用计算机画出，并且能够标明表面是由哪些残基的何种原子所组成的。为了细致了解分子表面形貌的变化情况，还可作出等高线图(图 5-30—32)。



图 5-30　胰岛素单体分子 I 逆 Z 轴方向观察到的分子表面形貌示意图(数值意义见图 5-32)

0:-15.3	5:-0.5
1:-12.3	6: 2.4
2:-9.4	7: 5.4
3:-6.4	8: 8.3
4:-3.5	9: 11.3



图 5-31　胰岛素单体分子 II 逆 Z 轴方向观察到分子表面形貌示意图

0:-22.1	5:-9.4
1:-19.6	6:-6.9
2:-17.0	7:-4.3
3:-14.5	8:-1.9
4:-12.0	9: 0.9

图 5-31 是逆 Z 轴方向观察到的胰岛素分子的 RV-1.7 表面"形貌"示意图。从图中可知，对胰岛素分子来讲，看不出其表面有鲜明的"洞穴"或"小峰"。

氢键网络和微环境分析是蛋白质分子结构研究的又一重要方面。

氢键网络协同疏水作用，是稳定蛋白质分子高级结构的重要因素。越来越多的事实表明，螺旋氢键的作用主要是结构性的，而非螺旋氢键的作用既有结构性的又有功能性的。蛋白质分子的活性部位，往往有复杂的氢键网络，这些氢键网络不仅是维持这一部位特征所必需的，而且作

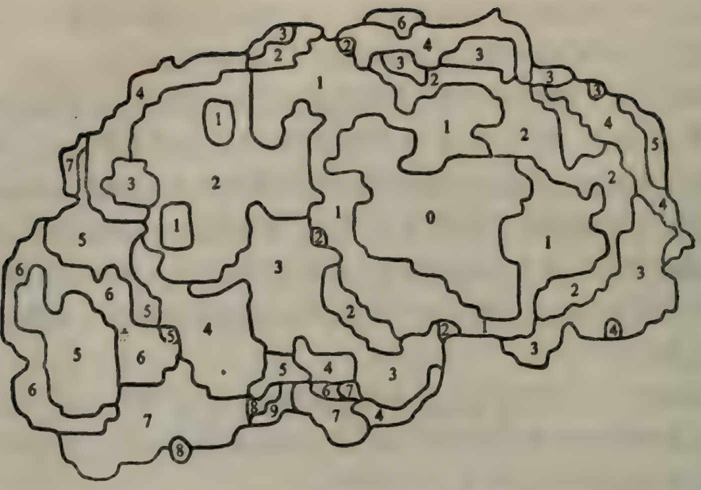


图 5-32 胰岛素双体逆 Z 轴方向观察到的分子表面形貌示意图(图中阿拉伯数字表示所在区域地势的高度，其值如下：0:-22.1；1:-18.5；2:-15.0；3:-11.4；4:-7.8；5:-4.3；6:-0.7；7:2.9；8:6.4；9:10.0)

为一种次级相互作用，氢键能使这一局部构象有相当的挠性，提供了分子反应的基础。

微环境，特别是活性部位的微环境，是研究分子反应所不可少的。蛋白质分子活性部位的微环境通常是复杂的，有其几何、物理和化学的特征。实际上这些特征的加强和削弱以至丧失，对结构、功能的影响以及经化学修饰后分子性能的预测等等，都是和微环境的研究分不开的。

我们建立了氢键网络和微环境搜索程序。图 5-33 是胰岛素二聚体中氢键网络的情况。

由于氢键的存在强烈地依赖于氢键键长和键角，所以在最终判断时，我们略去了所有那些键长大于 0.32nm，键角小于 120°的可能氢键。

从图中可以看到，胰岛素分子主链内富含氢键，它们主要是肽键自身的 NH 和 CO 间的键合。A、B 两链之间和二聚体两个单体之间都有若干氢键相连。

微环境搜索表明，A 链的情况要比 B 链的复杂。A 链的 A_{21}-天冬氨酸是微环境最复杂的一个残基。它的羰基氧和羧基氧在 0.5nm 范围内就包括了 3 个残基的 33 个原子。正是它的微环境的复杂性，使得 A_{21}-天冬氨酸不仅在保持胰岛素三维结构的稳定性方面，而且在其发挥胰岛素的功能方面都占有重要的地位。

5.4.1.2 胰岛素分子静电势的计算

在第 2 章中，我们已经介绍了分子静电势方法。若采用 CNDO/2 中的基本近似 $\Phi_\mu \Phi_\nu = \delta_{\mu\nu}$，那么在 CNDO/2 方法中，分子静电势 $V(\vec{r})$ 可表示为

$$V(\vec{r}_1)=\sum_a (Z_a/r_{1a}) - \sum_\mu P_{\mu\mu}\int \frac{\Phi_\mu^2(\vec{r}_2)}{r_{12}} d\vec{r}_2$$

式中 Z_a 为原子 a 的价层电子数，r_{1a} 和 r_{12} 分别是原子核 a 和点 $\vec{r}_2$ 至 $\vec{r}_1$ 的距离。

若分子表面点取 RV＝1.7，则可证明

$$V(\vec{r}_1)=\sum_a Q_a/r_{1a}$$

此时 Q_a 是分子中原子 a 所携带的净电荷数。

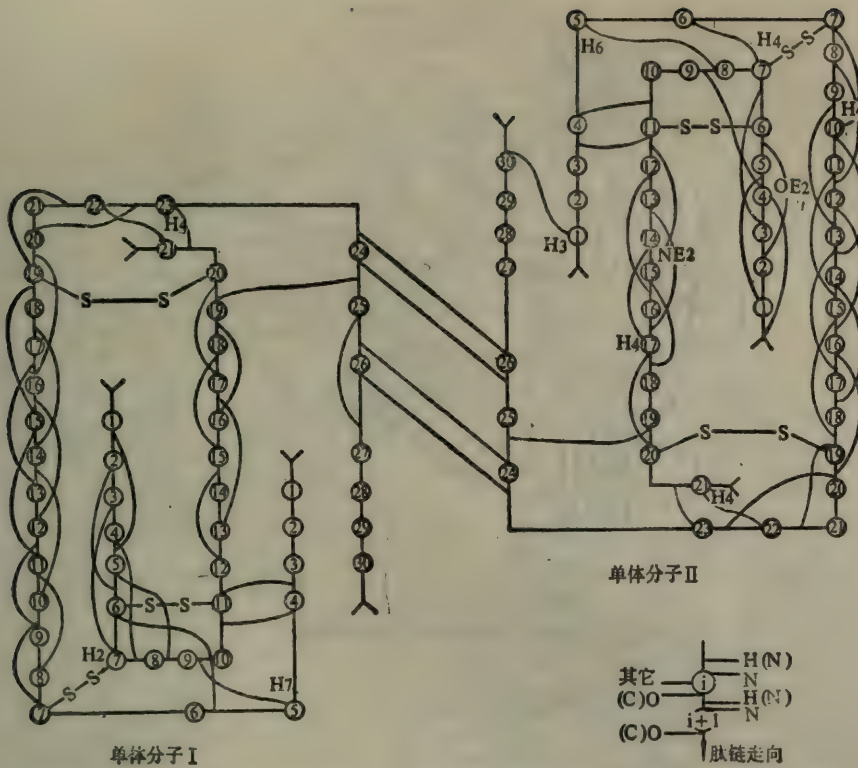


图 5-33 胰岛素二聚体中氢键网络概况示意



图 5-34 分子Ⅰ ZX-DOWN 面的静电势

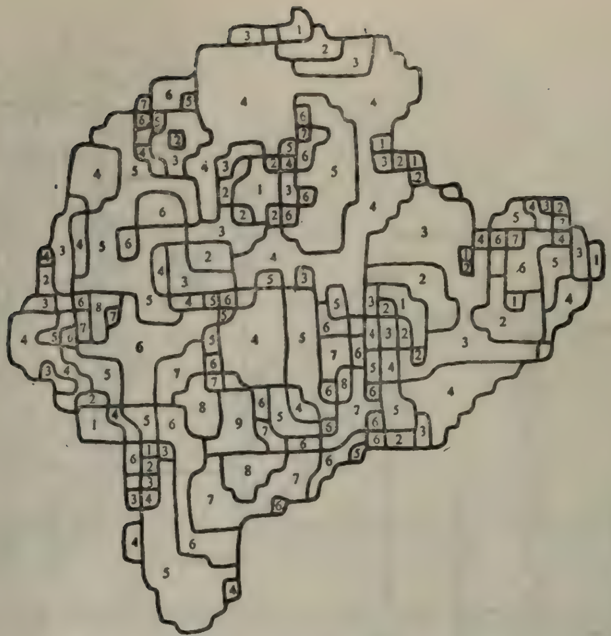


图 5-35 分子 II ZX-DOWN 面的静电势



图 5-36 双体 ZX-DOWN 面的静电势



图 5-37　分子 I *ZX*-UP 面的静电势



图 5-38　分子 II *ZX*-UP 面的静电势

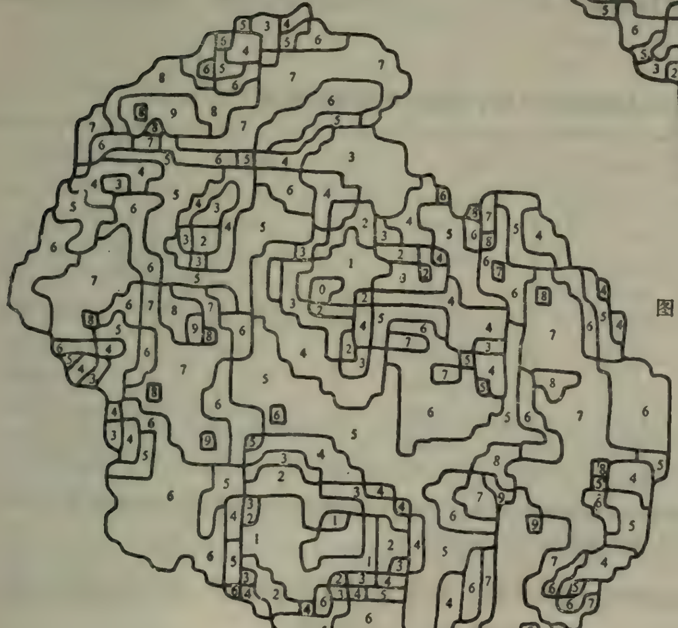


图 5-39　双体 *ZX*-UP 面的静电势

作者曾采用肽分割法，求出了胰岛素分子中的所有 Q_a，然后按上式便可得所需结果。全部数值计算结果在此只能采用图示法。我们给出其中的几幅(图 5-34—39)。

图5-34、35、36分别是沿 x 轴方向观察胰岛素分子 I、胰岛素分子 II 及胰岛素双体分子时所见到的分子表面静电势。该面记为 ZX-DOWN 面。

图5-37、38、39分别是逆 x 轴方向所见的分子表面静电势，该面记为 ZX-UP 面。

图中阿拉伯数字表示电势的大小，0 表示最负值，9 表示最大正值。具体数值范围按下述方法确定：电势最负值的绝对值加上电势的最大正值，将其分成十个等分，每一个等分用一个阿拉伯数字表示。最大正区用 9 表示，最小负区用 0 表示。

从图中可以清楚地看到，胰岛素分子，无论是分子 I 还是分子 II，它们的分子表面静电势都呈现鲜明的偶极型分布。负端为 A 链羧端，特别是A_{20}-半胱氨酸和A_{21}-门冬酰氨所在区域。正端为 B 链的 B_8-甘氨酸，B_9-丝氨酸和B_{10}-组氨酸所在区域。

分子 I 最负电势为 −188.28kJ/mol，最正电势为 +159 kJ/mol。分子 II 最负电势为 −167.36kJ/mol，最正电势为 +150.62kJ/mol。

两个单体分子的B_{23}-甘氨酸，B_{24}-苯丙氨酸所处的电势都落在 ±41.84kJ/mol 之内，是分子中最电中性的区域。

分子 I 和分子 II 的分子表面静电势的显著差别主要表现在 A 链的氨端。胰岛素单体分子I在A_4-谷氨酸及其邻区有一个较强的正势区，使得单体分子 I 的电势有类偶极型分布的特点。但是单体分子 II 在 A_4-谷氨酸处没有这种特点。相反，它具有中等强度的负势。从图5-34 和图5-35及表5-11、12可以看到这种差别。

表 5-11　分子 I 和分子 IIA 链氨端若干原子所在区域的电势深度 (kJ/mol)

原　子	分子I	分子II	原　子	分子I	分子II
A3CG1*	94.56	35.48	A5CA	10.54	−34.52
A3H5	103.34	−43.01	A5CG	−77.19	36.44
A4CB	59.29	54.81	A5CD	−21.59	−36.94
A4CG	96.98	61.76	A5OE1	−85.98	−118.37
A4OE1	8.41	−111.21	A5NE2	63.51	−34.69
A4H3	50.84	45.65	A5H2	−31.59	−31.05
A4H4	31.46	32.18	A5H5	35.52	18.20
A4H5	106.40	19.29	A5H8	42.51	−44.56
A4H6	80.29	50.71			

胰岛素分子A链第三个氨基酸上的CG1原子，原子名称按IUPAC规则命名。

另外，在 B_{27}-精氨酸处，电势深度有较大的不同。单体分子 I 在该处的分子静电势落在±41.84kJ/mol之内，而单体分子 II 都高达±133.89kJ/mol。

从表5-12可以看到，在 B 链的B_{25}-苯丙氨酸、B_{26}-酪氨酸所在区域，单体分子 I 要比单体分子 II 温和些。

从图 5-34、35，图 5-37 及图 5-38、39 的比较中，可以看到单体形成双体后，静电势的变化趋

表 5-12 单体和双体 B 链若干原子所在区域的电势深度(kJ/mol)

原子	单体中		双体中	
	分子 I	分子 II	分子 I	分子 II
B25CB	−24.48	−41.80	−21.97	
B25CG	−30.17	−31.51	−31.25	−40.92
B25CD1	−80.37	−15.02	−66.82	−28.82
B25CE1	−104.81	−20.92	−87.99	−32.93
B25CZ	−50.84	−91.13	−40.17	−91.76
B25CE2	−22.22	−99.08	−36.98	−96.57
B25CD2	−25.82	−99.45	−29.71	−98.20
B25H2	17.07	49.12		
B25H3	−5.61	−66.02	−6.99	
B25H4	−70.42	40.92	−59.16	
B25H6	−89.24	18.58	−72.55	
B25H7	−19.46	−70.71	−34.69	−72.97
B25H8	8.03	−163.34	−82.84	−160.21
B26CB	62.80	77.74		
B26CZ	102.42	106.11		
B26OH	86.06	82.89	88.28	65.06
B26O	−19.29	45.06		
B26H1	46.19	60.25		
B26H3	52.97	60.80		
B26H7	117.15	120.08	137.70	85.19
B27CA	28.45	136.19	7.15	
B27CB	32.13	118.37	−14.90	68.12
B27OG1	−29.54	14.90	−30.84	−26.82
B27CG2	41.84	93.14	−17.95	−15.02
B27H2	38.62	100.04	31.13	
B27H3	27.99	98.49	−17.23	55.02
B27H5	18.54	42.47		
B27H6	31.59	76.44	15.69	36.94
B27H7	−19.96	12.47	−21.34	−11.59

势。

计算结果表明,单体的正、负端在双体中得到了保证,原来的最负势和最强正势在集聚后仍分别为最负势和最强正势,但电势深度略有变化(图5-36和图5-39)。

单体集聚后,在其结合部位,分子表面情况发生了很大的变化,原来暴露在表面的 B_{23}-甘氨酸和 B_{24}-苯丙氨酸在双体中都成为双体的疏水内核,但其分子表面静电势除单体分子 II 的B_{27}-精氨酸急剧下降外,其余变化不大(表5-12),这可能是由于单体 II 在该区域的表面构成在集聚过程中变化较大。

值得一提的是,表中所示原子电势深度是指由这些原子所形成的那部分分子表面的电势深度,反映的是分子整体的电势特征。

5.4.1.3 胰岛素分子相互作用能的计算

胰岛素分子间的相互作用是人们所关心的。它对探讨胰岛素A、B链的重组，胰岛素及其类似物与受体的相互作用以及它们在溶液中的行为有着重要的意义。

下面我们以两个胰岛素分子在不同距离和不同取向时的相互作用能的计算为例作一初步分析。

根据量子力学微扰理论，在零微分重叠近似的基础上，采用点电荷模型，分子A和分子B之间的相互作用能可以分解为静电能 $E_{el.}$、极化能 $E_{pol.}$、色散能 $E_{dis.}$和近程排斥能 E_{rep}。

$$E = E_{el.} + E_{pol.} + E_{dis.} + E_{rep}$$

其中

$$E_{el.} = -\frac{1}{2}\sum_{a=1}^{N_A}\sum_{b=1}^{N_B}\rho_a\rho_b/r_{ab}$$

$$E_{pol.} = -\frac{1}{2}\sum_{a=1}^{N_A}\sum_{b=1}^{N_B}(\alpha_a\rho_b^2+\alpha_b\rho_a^2)/r_{ab}^4$$

而 $E_{dis.}+E_{rep.}$通常表示成Buckinham势的形式：

$$E_{dis.}+E_{rep.}=K_aK_b[-A/Z^6+(1-\rho_a/N_a^{val})(1-\rho_b/N_b^{val})\cdot C\cdot e^{-\beta z}]$$

以上诸式中，ρ 表示分子中原子所携带的净电荷数，α 表示分子中原子的平均极化率，r 表示距离，N^{val}表示分子中原子的价电子数，其余均为常数。

在对胰岛素分子间相互作用能进行计算时，我们从胰岛素两聚体的晶体结构数据出发，先找出两个单体分子相距最近的三对点，求其中点坐标。在过该三点的平面上作一条垂线，作为分子间平移和转动的参考轴。固定分子I，沿参考轴平移、旋转分子II，就可以得到两个分子相距不同距离 d 和采取不同取向 θ 时的全坐标。

$\theta=0^\circ$同时 $d=0$表示两聚体取晶体结构数据，即一般认为的平衡点位置；$d>0$表示彼此远离；$d<0$表示比 $d=0$更为接近。θ 的取值范围为0°—360°。

有了根据上述方法所确定的两个胰岛素分子的全坐标，就能够计算分子间的相互作用并分析它随分子间距离 d 和分子间相对取向 θ 的关系。

表5-13给出了分子I和分子II在 $d>0.10$nm 的不同距离和不同取向 θ 时的静电能 $E_{el.}$和非键原子间的近程排斥能 $E_{rep.}$的部分计算结果。

从表5-13中可以看到，当 d 一定时，$E_{el.}$和 $E_{rep.}$ 随 θ 变化趋势大致相同。在30°—210°，两分子表现出强烈的静电互斥，即$E_{el.}+E_{rep.}\gg 0$，且往往在60°时达到最大。在210°—320°，$E_{el.}$和 $E_{rep.}$的量级基本相同，两分子间的静电作用能$E_{el.}+E_{rep.}$约为 10—10^2kJ/mol。在0°附近，上述二项之和迅速下降，$E_{el.}+E_{rep.}$恒为负值，且在0°达到最小。例如，$\theta=0^\circ$，$d=0.16$nm 和 $d=0.11$nm时，$E_{el.}+E_{rep.}$约为-20kJ/mol 和-26kJ/mol。因此，若仅考虑两个分子间的静电和近程互斥，就可以说：当两个胰岛素分子沿我们指定的那根参考轴相互接近时，它们的相对取向角度将落在0°附近。因为只有在这范围中，分子彼此靠扰时仅需克服很低的能障或者不需克服任何能障就能使整个体系进一步稳定。从图5-40的总相互作用能与 d、θ 的关系图中可以清楚地看到这一点。

*表 5-13 静电能$E_{el.}$和非键原子近程排斥能$E_{rep.}$(单位：kJ/mol)

	0°	60°	120°	210°	300°
0.11nm	−34.49	68.49	31.88	−41.29	−27.52
	8.32	$\sim 2\times10^4$	$\sim 2\times10^5$	245.99	40.58
0.13nm	−30.40	83.31	32.25	−39.96	−26.19
	5.32	$\sim 2\times10^4$	$\sim 2\times10^5$	161.07	22.60
0.16nm	−25.41	56.35	51.05	−24.46	−18.03
	2.63	$\sim 1\times10^4$	$\sim 5\times10^4$	118.56	18.94
0.30nm		51.25	22.63		
		$\sim 3\times10^3$	$\sim 2\times10^3$		

* 每项中第一行表示$E_{el.}$，第二行表示$E_{rep.}$

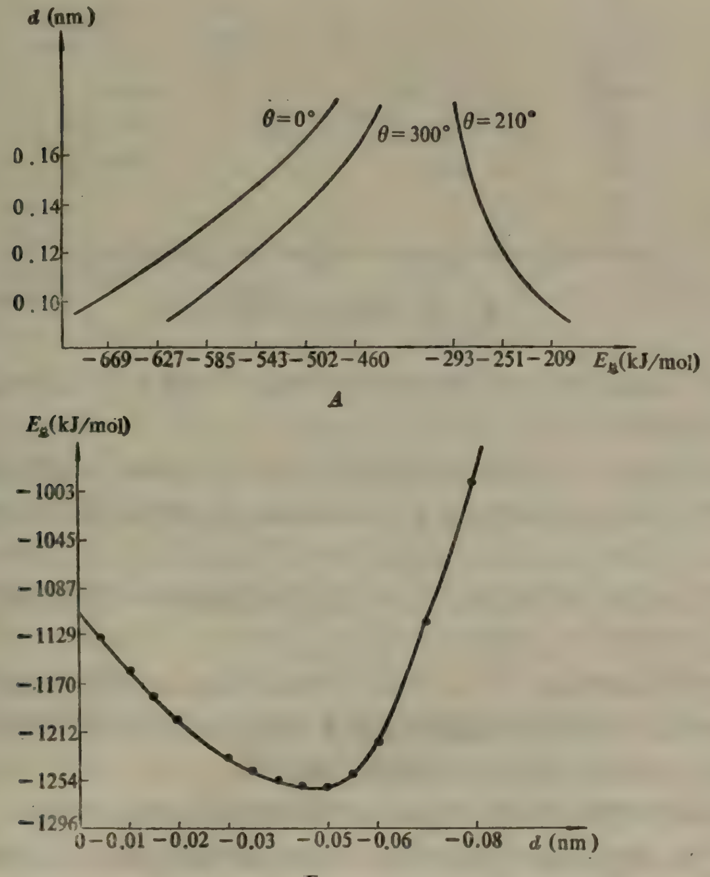


图 5-40 总相互作用能与 d 的关系

A. $d>0$时，$E_总$与θ、d的关系

B. $d<0$时，$E_总$与 d 的关系

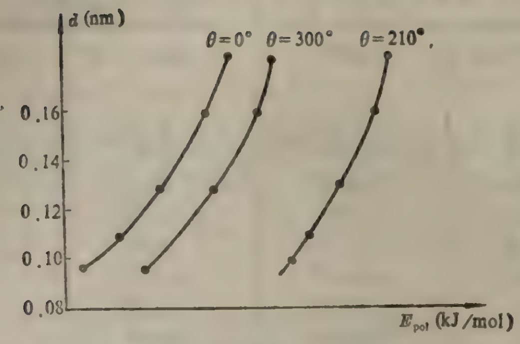


图 5-41 极化能 $E_{pol.}$ 与 d 的关系

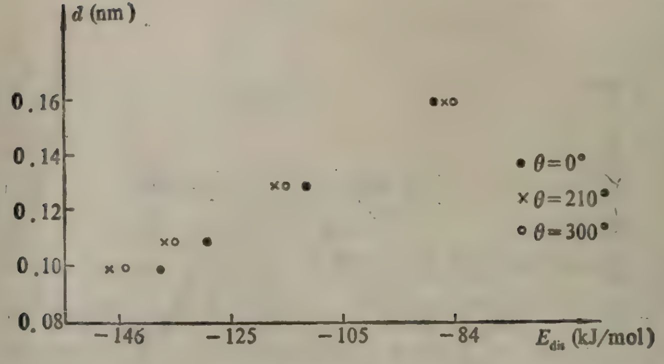


图 5-42 色散能 $E_{dis.}$ 与 d 的关系

图5-41和图5-42画出了0°、210°、300°时分子 I 和分子 II在 $d>0.10$nm时的极化能 $E_{pol.}$和色散能 $E_{dis.}$与 d 的关系。

从计算结果可以发现，当 d 固定时，色散能$E_{dis.}$ 随 θ 变化不大。d 愈大，这种变化愈小。这是因为色散能跟 Z 的负六次方成正比之故。当 θ 从300°变化到360°时，极化能 $E_{pol.}$是逐渐下降的。

从$E_{el.}$ 和 $E_{rep.}$ 有利于分子相互靠近、体系能量降低的区域中，不难发现，$E_{dis.}+E_{pol.}$ 也恰在0°时达到最小。

综上所述，我们可以得出如下结论：不论分子相距多远，从相互作用能的观点来看，它们将沿着某一方向相互靠近。在我们所选定的参考轴、参照系中，分子 II 是沿 θ 为零的方向靠近分子 I 直至平衡位置的。在此整个过程中，体系能量不断下降。

为寻求平衡点，我们还计算了 $\theta=0$，$d<0$将近 20点相互作用能计算值。

从图5-40中可以看到，当 $\theta=0$，但$d<0$时，体系的总相互作用能开始是随$|d|$的增大很有规律地降低，当 $d=-0.05$nm时，总相互作用能为-1261.69kJ/mol。当$|d|>0.05$nm时，总相互作用能迅速上升，当 $d=-0.08$nm时，总相互作用能为-999kJ/mol左右。当 $d=-0.10$nm时，总相互作用能就成正值，即两分子互斥了。由此可见，我们的理论计算得出的稳定平衡点位置是

$\theta=0°, d=-0.05$nm。 考虑到我们的计算没有计及晶体中多个胰岛素分子间的相互作用，特别是没有计及环境水的影响，再加上计算模型的近似性，可以认为，我们的计算和晶体结构测定所得出的关于稳定平衡点的结论是相当吻合的。

表5-14给出了 $\theta=0, d$ 取不同值时相互作用能的组成状况。从表中可以看到，极化能 $E_{pol.}$ 总是占总相互作用能中的绝大部分，色散能 $E_{dis.}$ 次之，近程排斥能最小。这和生物大分子在相互接近过程中相互作用的本质是一致的。

下面着重讨论 $\theta=0°$ 时，胰岛素分子中各氨基酸残基在分子间相互作用能的贡献和意义。

表5-14列出了 $\theta=0°, d$ 分别为0.05nm、0.11nm、0.13nm、和0.16nm 时，分子 I、分子 I 中A链羧端的 A_{19}、A_{21}以及 B链的B_{12}、B_{16}、B_{21}、B_{22}、B_{24}、B_{25}和B_{26}与分子 II 及分子 II 中A链、B 链的上述残基的相互作用能。

表 5-14 相互作用能的组成 ($\theta=0°$)(单位：kJ/mol)

	E_{el}	E_{pol}	E_{dis}	E_{rep}	$E_{总}$
−0.05nm	−108.38	−1046.94	−606.78	502.56	−1261.69
0nm	−73.71	−798.27	−360.71	123.16	−1109.54
0.11nm	−34.49	−497.00	−130.86	8.31	−654.00
0.13nm	−30.40	−461.64	−110.89	5.32	−597.61
0.16nm	−25.41	−415.27	−85.05	2.63	−525.13

从表中可以看到：分子 I(或分子 II)的B_{12}、B_{13}、B_{16}、B_{24}、B_{25}、B_{26}中和分子 II(或分子 I) 的相互作用极其强烈，其中一分子中的B_{26}与另一分子的相互作用能在 d取0.16、0.13、0.11和0.05nm时高达−53kJ/mol、−63.5kJ/mol、−71kJ/mol和−103.6kJ/mol。两个分子中上述氨基酸残基间的作用也很强烈。例如，一个分子的B_{26}与另一个分子的B_{24}间的相互作用在 d 取0.16、0.13、0.11和0.05nm时可达到−13.5kJ/mol、−17.2kJ/mol、−20.6kJ/mol 和−35.3 kJ/mol。由于这些残基又是非极性的(B_{13}除外)，因此，从该表就可以看出：在胰岛素分子的集聚过程中，上述残基是两聚体的结合部位，它们之间的相互作用对两聚体的稳定有着重要的作用。晶体结构测定和胰岛素结构数据的计算机处理也均指明了B_{12}、B_{16}、B_{24}、B_{26}确实是分子 I 和分子 II 的结合部位，特别是其中一分子B_{24}、B_{26}与另一分子的B'_{26}、B'_{24}间的两对氢键在两聚体的稳定化中起着极为重要的作用。

我们的计算不仅定性而且定量地对这些残基在结合部位中的作用给出了评价。

其次，从表中还可以看到 A 链中 A_{21}的地位。在分子间相互作用能中，A_{21}不仅是 A 链中最为突出的一个残基，在整个分子中也是较为显著的。这种现象跟A_{21}的极性和复杂的微环境有着密切的关系。实验表明：A_{21}对胰岛素的分子构象和生物活力是极其敏感的。我们的计算对实验的理解是有启发意义的。值得一提的是在不同种属的胰岛素分子中，仅作保守替换(恒为酸性)的B_{13}比具有与 A_{21}相类似的性质。

最后，从表5-15的非对称性可以断定分子 I 和分子 II 的构象是不同的。从表格非对称性的

表 5-15　两分子若干残基间的相互作用能(kJ/mol)

		A_{10}	A_{21}	B_{12}	B_{13}	B_{16}	B_{21}	B_{22}	B_{24}	B_{25}	B_{26}	分子II
0.11nm	A_{10}	−0.130	−0.385	−0.205	−0.150	−0.176	−0.217	−0.222	−0.631	−0.874	−0.552	−6.830
0.13nm		−0.125	−0.364	−0.196	−0.146	−0.167	−0.213	−0.213	−0.589	−0.815	−0.514	−6.508
0.16nm		−0.117	−0.334	−0.184	−0.138	−0.159	−0.201	−0.196	−0.531	−0.740	−0.460	−6.061
0.05nm		−0.150	−0.456	−0.230	−0.163	−0.192	−0.247	−0.255	−0.790	−1.083	−0.694	−7.938
	A_{21}	−0.510	0.071	−0.276	−0.079	−0.159	+0.063	−0.079	−0.828	−1.714	−7.047	−27.709
		−0.485	0.088	−0.268	−0.075	−0.155	+0.063	−0.075	−0.807	−1.584	−6.387	−25.423
		−0.447	0.113	−0.255	−0.075	−0.150	0.063	−0.071	−0.773	−1.409	−5.530	−22.518
		−0.600	0.125	−0.305	−0.088	−1.705	0.063	−0.088	−0.890	−2.224	−9.530	−36.930
	B_{12}	−0.209	−0.431	−8.368	−4.999	−4.698	−0.263	−0.226	−4.356	−0.719	−1.154	−36.604
		−0.201	−0.418	−7.248	−4.565	−4.092	−0.259	−0.217	−3.900	−0.690	−1.099	−33.210
		−0.188	−0.401	−5.601	−3.992	−3.352	−0.251	−0.203	−3.185	−0.652	−1.020	−28.125
		−0.234	−0.460	−12.866	−6.684	−7.230	−0.284	−0.255	−5.484	−0.803	−1.325	−49.044
	B_{13}	−0.166	0.588	−4.427	−25.352	−1.542	−0.201	−0.146	−0.869	−0.268	−0.949	−53.395
		−0.159	0.588	−4.092	−22.095	−1.480	−0.196	−0.142	−0.436	−0.259	−0.915	−48.170
		−0.155	0.063	−3.645	−17.945	−1.392	−0.192	−0.142	−0.786	−0.251	−0.869	−41.420
		−0.184	0.053	−5.668	−37.143	−1.747	−0.209	−0.150	−0.982	−0.284	−1.053	−72.469
	B_{16}	−0.209	−0.155	−5.137	−0.983	−0.268	−0.063	−0.079	−0.631	−0.623	−8.034	−44.450
		−0.201	−0.150	−4.131	−0.936	−0.259	−0.063	−0.079	−0.602	−0.602	−7.127	−40.178
		−0.192	−0.146	−3.658	−0.869	−0.242	−0.063	−0.075	−0.564	−0.573	−6.015	−34.811
		−0.234	−0.163	−8.021	−1.150	−0.301	−0.069	−0.088	−0.720	−0.694	−11.106	−41.517
	B_{21}	−0.205	−0.025	−0.506	−0.309	−0.142	−0.054	−0.075	−0.297	−0.330	−3.256	−34.238
		−0.196	−0.021	−0.493	−0.305	−0.138	−0.054	−0.075	−0.293	−0.320	−3.031	−31.847
		−0.192	−0.021	−0.472	−0.297	−0.138	−0.050	−0.071	−0.284	−0.314	−2.721	−25.435
		−0.230	−0.029	−0.552	−0.326	−0.146	−0.054	−0.079	−0.314	−0.347	−4.042	−42.954
	B_{22}	−0.234	−0.050	−0.209	−0.050	−0.067	−0.029	−0.054	−0.389	−0.911	−2.521	−22.203
		−0.226	−0.050	−0.201	−0.050	−0.067	−0.029	−0.054	−0.372	−0.869	−2.307	−20.511
		−0.209	−0.046	−0.192	−0.046	−0.067	−0.025	−0.050	−0.351	−0.819	−2.023	−18.271
		−0.272	−0.054	−0.234	−0.051	−0.071	−0.029	−0.054	−0.431	−1.032	−3.318	−28.570
	B_{24}	−0.510	−0.401	−3.102	−0.819	−0.535	−0.334	−0.355	−5.948	−9.075	−19.533	−55.623
		−0.472	−0.385	−2.725	−0.786	−0.514	−0.326	−0.342	−5.242	−8.030	−16.206	−49.709
		−0.426	−0.359	−2.274	−0.736	−0.481	−0.314	−0.326	−4.364	−6.776	−12.490	−42.030
		−0.640	−0.460	−4.744	−0.936	−0.614	−0.355	−0.337	−8.891	−13.439	−33.093	−82.166
	B_{25}	−1.062	−6.157	−0.702	−0.418	−0.564	−0.489	−0.911	−10.308	−9.614	−4.121	−43.071
		−0.986	−5.551	−0.677	−0.410	−0.539	−0.485	−0.874	−9.213	−8.318	−3.783	−39.129
		−0.890	−4.769	−0.640	−0.401	−0.506	−0.472	−0.815	−7.833	−6.692	−3.678	−34.038
		−1.325	−8.444	−0.786	−0.443	−0.640	−0.506	−1.035	−14.730	−14.128	−5.396	−57.680
	B_{26}	−0.564	−8.406	−1.137	−1.087	−6.048	−2.238	−2.546	−20.612	−3.223	−1.601	−71.478
		−0.518	−7.654	−1.078	−1.041	−5.396	−2.065	−2.345	−17.251	−2.918	−1.513	−63.452
		−0.468	−6.646	−0.995	−0.982	−4.636	−1.843	−2.073	−13.276	−2.529	−1.388	−53.496
		−0.715	−11.140	−1.329	−1.225	−8.519	−2.780	−3.277	−35.367	−4.448	−1.889	−104.053
	分子II	−6.071	−30.718	−34.080	−59.059	−43.639	−24.320	−14.981	−58.382	−33.269	−72.569	−654.003
		−6.379	−28.219	−30.890	−53.262	−39.284	−22.526	−13.940	−51.849	−30.092	−64.397	−597.615
		−5.931	−24.946	−26.836	−45.746	−33.766	−20.068	−12.565	−43.700	−26.037	−54.319	−525.733
		−7.821	−40.487	−46.703	−80.118	−60.740	−31.751	−18.885	−84.787	−45.240	−105.160	−871.948

大小可以初步估计构象偏差的程度。

5.4.1.4 胰岛素分子主链的CNDO/2计算

要对整个生物大分子进行哪怕是最简单的量子化学计算，例如 EHMO 法，至少在目前还是不可能的。寻找能对生物大分子进行量子化学计算的方法和技巧是人们所关注的一个课题。

分割法，即将大分子分成几块碎片，是比较可行的方法。例如，Pullman，A.等在对核酸进行量子化学计算时，把碱基、糖和磷酸根作为最基本的计算单元，采用分割法，使对核酸的计算降级为对上述基本单元的计算。在对蛋白质进行量子化学计算时，则可以肽、残基为基本计算单元，还原蛋白质的量子化学计算为对各肽、残基的计算。木瓜蛋白酶和胰岛素的量子化学研究采用的便是分割法。

分割法的最主要缺陷是不能全面反映分子的二级结构。而大分子的二级结构，即主链构象，恰恰是决定分子三维折叠和决定分子生物功能最重要的结构层次。

为此，我们提出了蛋白质分子非肽分割的量子化学计算方案，如图 5-43 所示，其中(*A*)是肽分割示意，(*B*)是我们提出的非肽分割示意。

非肽分割法在计算时考虑了所有主链原子的位置，自然地包含了主链的全部构象信息，因此能比较满意地表现出蛋白质分子三级结构对计算的影响，这是非肽分割法最突出的优点之一。对半经验的量子化学计算方法，非肽分割法尤为适用。侧链的构象可以通过以肽为基本单位的计算加以补充，甚至可以采用标准数据予以补充。

从化学或生物学意义上来说，非肽分割方案显然优于肽分割法。

该思想可自然推广至其它生物高分子的计算。

该方法还可用于链内小环或链间桥键的计算。

在对胰岛素分子进行 CNDO/2 计算时，根据胰岛素的结构特点和计算机容量，我们将胰岛素A 链的主链分成三段，B链之主链也分成三段，具体情况如下：

各肽段结构

肽　　段	原子数	价电子数	价轨道数	结构特点
A_1—A_9	65	200	173	螺旋结构
A_{10}—A_{18}	65	200	173	A_{12}～A_{17} 为 3_{10} 螺旋
A_{19}—A_{21}	24	74	63	
B_1—B_{10}	72	222	192	B_1～B_8 为松弛伸展
B_{11}—B_{20}	72	222	192	B_9～B_{20} 为 α- 螺旋
B_{21}—B_{30}	73	228	196	B_{24}～B_{30} 为自然折迭
A_6—A_{11}	46	146	124	双硫键

此外，我们还计算了包括A_6—A_{11}间双硫键在内的A链小环。

5.4.2 木瓜蛋白酶的量子化学研究

木瓜蛋白酶是一种半胱氨酸蛋白酶，分子量 23 000。它是由 212 个氨基酸联结成的一条多肽链。Drenth 和其合作者于1976年用 X-射线衍射法测定了木瓜蛋白酶的晶体结构和原子坐标。

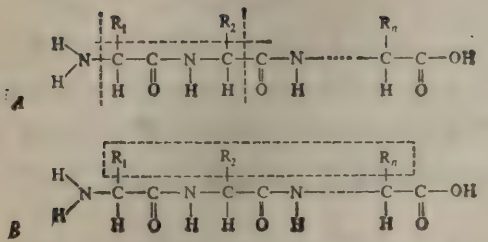


图 5-43

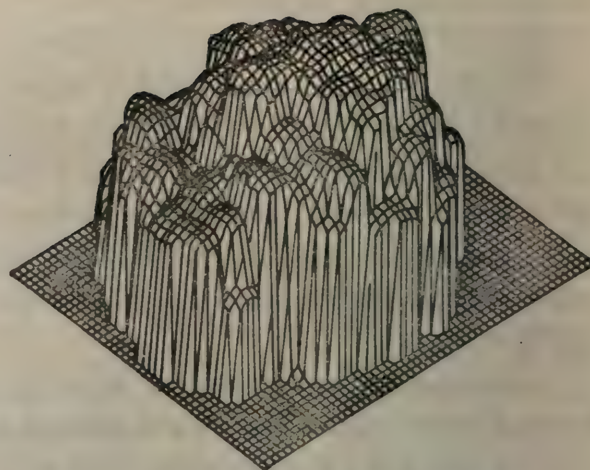


图 5-44 木瓜蛋白酶分子在XZ平面方向上朝Y面的分子表面外壳

图 5-44 画出了木瓜蛋白酶分子在 XZ 平面方向上朝 Y 面的分子表面外壳。

图 5-45 画出了木瓜蛋白酶分子在 XY 平面方向上朝 Z 面的分子表面外壳。

图 5-46 是活性中心区域在XY平面方向上朝Z面的分子表面外壳。

从这些图中，可以发现木瓜蛋白酶分子，尤其是其活性中心区域有以下几何特征：

1）整个木瓜蛋白酶分子由二部分或二翼构成，从图 5-47 仅由CA原子构成的蛋白质分子主链结构图可以看得更清楚些。其中 11—110，207—212 号氨基酸组成一翼，1—10，111—207 号氨基酸组成另一翼。从图 5-44 和图 5-45 也能看到这一特征。整个分子形状好像一个动物大脑，由二个半球组成，在此二半球之间有一条相当深的沟缝。

2）活性中心区位于这条沟缝的深处底部，其中半胱氨酸 25 和组氨酸 159 各位于一边，隔过

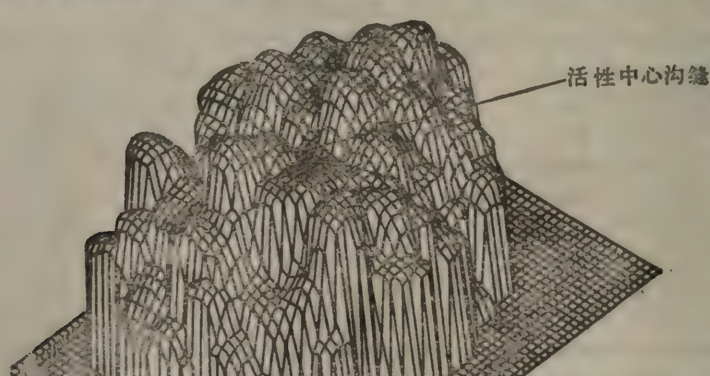


图 5-45 木瓜蛋白酶分子在XY平面方向上朝Z面的分子表面外壳

这条沟缝相向而对。很多蛋白酶，如溶菌酶、核糖核酸酶和枯草杆菌酶等都有这种活性中心在一道沟缝深处的几何特征。

3）在活性中心区，关键性的残基组氨酸159和半胱氨酸25都暴露在分子表面。若我们从Z方向看木瓜蛋白酶的XY平面时，就可以看到沟缝中的组氨酸159和半胱氨酸25，它们不被其它氨基酸肽链或侧链所遮盖。这样，底物分子才可以借助于嵌插(intercalation)作用而插入沟缝与之反应。这种特殊的分子几何学是生物大分子反应、分子识别时所必需之特征。这就理解了为什么如果有关残基如半胱氨酸25、组氨酸159或天冬酰胺175的相对位置有所变动时，或者其它侧链的构象稍有变动时，木瓜蛋白酶就可能改变甚至失去其反应活性。



图 5-46 木瓜蛋白酶分子活性中心区域在 XY 平面方向上朝Z面的分子表面外壳。

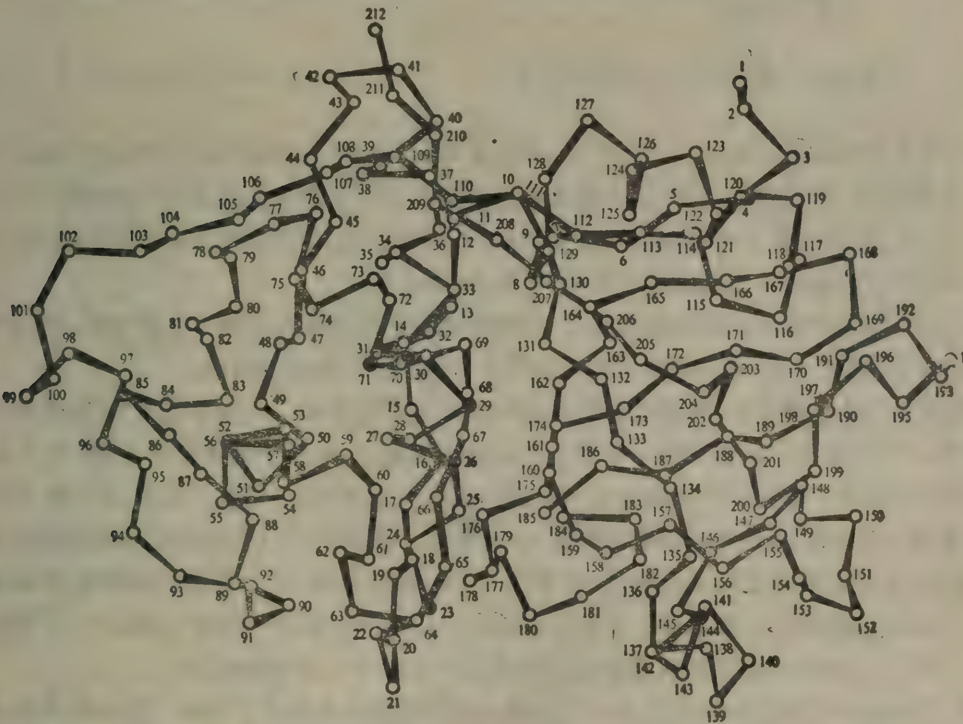


图 5-47 木瓜蛋白酶分子的主链(CA原子)结构

木瓜蛋白酶分子的活性中心主要由半胱氨酸25、组氨酸159和天冬酰胺175组成(图5-48)。这是一个"电荷转接系统"(charge relay system)，Drenth认为其催化机制可简述如下：

在活性中心中，两个主要残基侧链半胱氨酸25和组氨酸159之间发生质子转移，形成离子对I_m^+……S^-。硫原子位于组氨酸159咪唑环的同一平面。离子对的形成将有利于在咪唑环和天

冬酰胺 175 之间形成氢键。当一种肽型底物与酶结合时，易断裂肽的羰基 C═O 和谷氨酰胺 19 的氨基之间所形成的氢键。羰基的极化有利于硫原子进攻碳原子而形成四面体中介物。在该结构中，易断裂肽的NH基指向天冬氨酸 158 的 C═O，而 N 的孤对电子指向咪唑环。从初始位置旋转 30°，环就能把质子给予NH.肽键就断裂，形成一个胺和一个共价络合物(酰基酶)。然后发生脱酰基作用。水分子用一种类似但方向相反的方向把质子给予咪唑环，形成第二个四面体中介物，断裂C—S键，形成一种酸式产物而使酶恢复到活性状态(见图 5-49)。

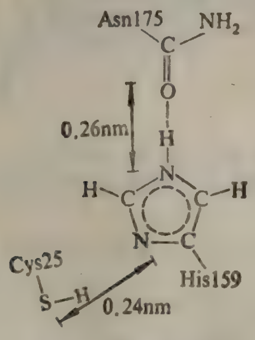


图 5-48 木瓜蛋白酶分子的活性中心，“电荷转接系统”。

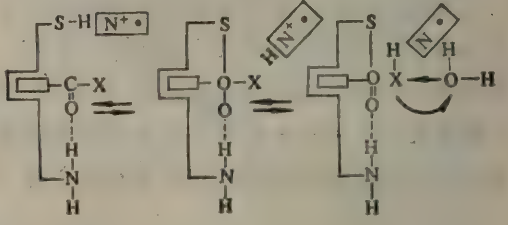


图 5-49 木瓜蛋白酶的催化机理

在对木瓜蛋白酶进行量子化学研究时，我们采用了 Pullman 等发展的一套方法，即首先把蛋白质分子(或核酸)分成许多次单元，也就是主链中的各个肽单位和各种侧链，对这些次单元作量子化学从头计算，进行电子密度的多极多中心展开。最后迭加各个次单元的静电势便可以得到整个分子的分子静电势。这种方案我们称为“肽分割”方案。

下面我们首先讨论木瓜蛋白酶的组分、肽基和各种不同类型侧链的分子静电势，然后考察各种肽基和侧链对整个木瓜蛋白酶分子和其活性中心的静电势分布的影响。

蛋白质分子是一种多肽，肽基是蛋白质结构中最基本的单元。在肽单元中.酰胺键是平面状的，即羰基 C═O 和 N—H 键在同一平面上。在这个平面上，氧原子带部分负电荷(图 5-50)，氢原子带部分正电荷。每个肽单元就变成一个偶极矩约为 3.5D的小偶极。木瓜蛋白酶中丙氨酸 136—丙氨酸 137 间肽基的分子静电势如图 5-51 所示。从图中可以看到明显的偶极静电势分布特点。显然，在蛋白质分子的某个部位，若存在一个α-螺旋结构，且α-螺旋结构中所有羰基键都指向同一方向(螺旋的轴)，那么每一个小偶极就会合成一个整体的大偶极，而形成一个较强的静电势场。如此，α-螺旋在蛋白质折叠和反应中常常会起着十分重要的作用。

木瓜蛋白酶分子经过分割产生 212 个肽基。这些肽基的构象各不相同。用标准构象角计算程序，算出了这些肽基中N-CA和CA-C间的二面角，并作构象图(图5-52)。根据构象图将这212个肽基分成十组，分别取 16、52、61、86、90、109、136、182、201 和 202 号肽基作为每一组的代表点。对这些代表点进行从头计算，这样便完成了 212 个肽基单元的计算工作。

非极性侧链　在氨基酸中，非极性侧链是指那些带脂肪烃链的甘氨酸、缬氨酸、亮氨酸、异亮氨酸和脯氨酸、一个芳香环侧链苯丙氨酸和一个甲硫氨酸。

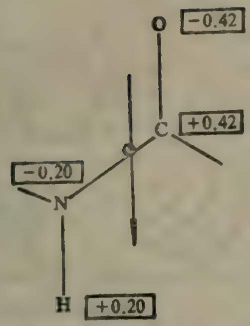


图 5-50　单个肽单元的偶极，偶极矩=3.7D。(24)

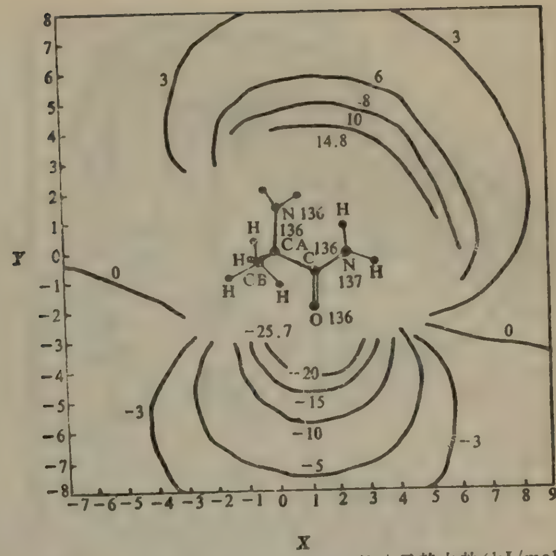


图 5-51　肽单元Ala136—Ala137 的分子静电势(kJ/mol)
1kcal=4.18kJ

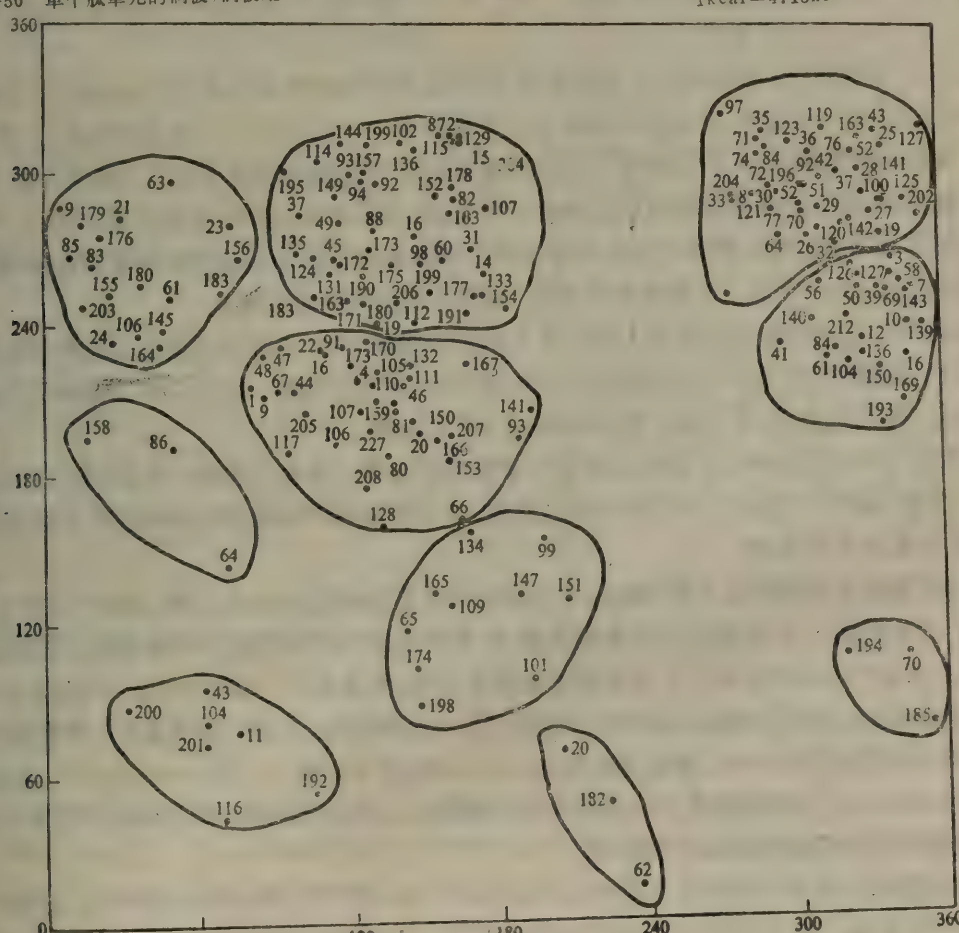


图 5-52　木瓜蛋白酶主链肽基的构象角图

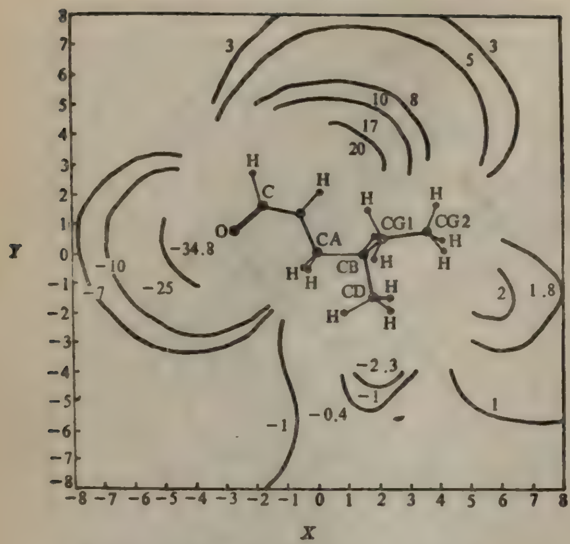


图 5-53 异亮氨酸 Ile 34 残基的分子静电势 (kcal/mol)

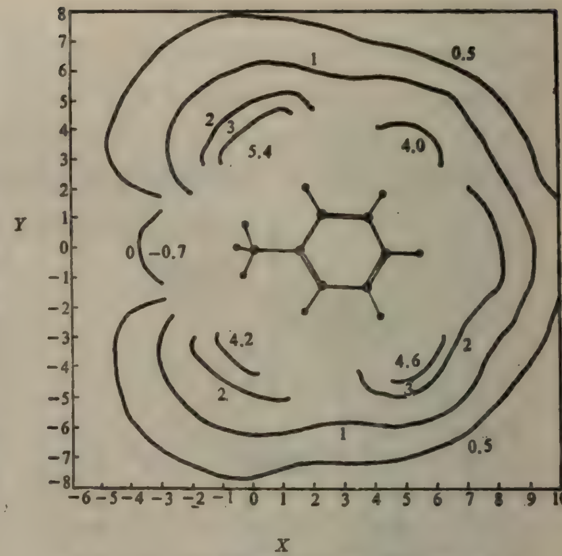


图 5-54 苯丙氨酸 Phe 28 侧链的分子静电势 (kcal/mol)

图 5-53 和图 5-54 分别画出了异亮氨酸 34 残基和苯丙氨酸侧链的分子静电势。从图 5-53 可知，对异亮氨酸 34 来说，仅仅在肽基部分由于羰基 C═O 键和 N—H 键间的偶极，才在空间有一个较重要的静电势分布，脂肪烃侧链产生的静电势相对来说是非常小的，一般在 ±21 kJ/mol 以下。图 5-54 是苯丙氨酸侧链所产生的分子静电势，与肽基部分相比也是非常小的。

丙氨酸、亮氨酸、脯氨酸和甲硫氨酸侧链的分子静电势分布与异亮氨酸和苯丙氨酸类似，一般也在 ±21 kJ/mol 以下。甘氨酸没有侧链，本身就是多肽链的一部分。

由此可见，非极性侧链在静电上即化学上是不活泼的，因而也是憎水性的。众所周知，在溶液中，非极性侧链往往分布在蛋白质分子的内部，很少在分子表面，这就是蛋白质分子结构所谓的“非极性（憎水结构）在里，极性（亲水结构）在外”的规则。

极性但不带电荷的侧链　极性但不带电荷的侧链包括二个含有羟基的氨基酸丝氨酸和苏氨酸，二个含酰胺基的氨基酸天冬酰胺和谷氨酰胺，二个带芳香环的氨基酸酪氨酸和色氨酸及一个带有硫的氨基酸半胱氨酸。

这一组带极性侧链的氨基酸在酶反应和催化中常常扮演重要角色。图 5-55、56、57 和 58 分别表示了天冬酰胺 175、丝氨酸 176、半胱氨酸 25 和色氨酸 177 侧链的分子静电势。从这些图中可以看出，侧链中的氧原子能产生相当强的负静电势场，羰基氧（−167 kJ/mol 左右）比羟基氧（−105 kJ/mol 左右）更强，这就导致这些侧链在化学上较活泼，并易于形成氢键。在许多蛋白质分子，如木瓜蛋白酶（papain）、胰蛋白酶（trypsin）、胰凝乳蛋白酶（α-chymotrypsin）、枯草杆菌蛋白酶（subtrlisin）、核糖核酸酶（ribonuclease）和弹性蛋白酶（elastase）等的活性中心中，天冬酰胺和丝氨酸等常常是活性中心组成部分。

半胱氨酸 Cys 侧链产生的分子静电势不大（图 5-60），但其巯基上的硫原子对于金属阳离子和其它反应物则是一个很强的电子授体，它还有与另一个半胱氨酸形成交联性二硫键的趋势，故

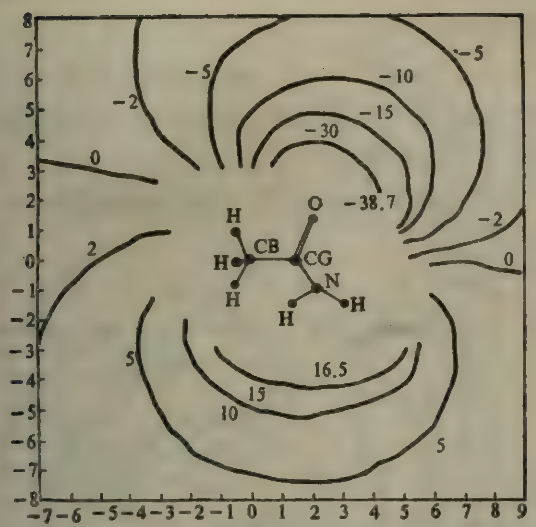


图 5-55 天冬酰胺 Asn 175 侧链的分子静电势
(kcal/mol)

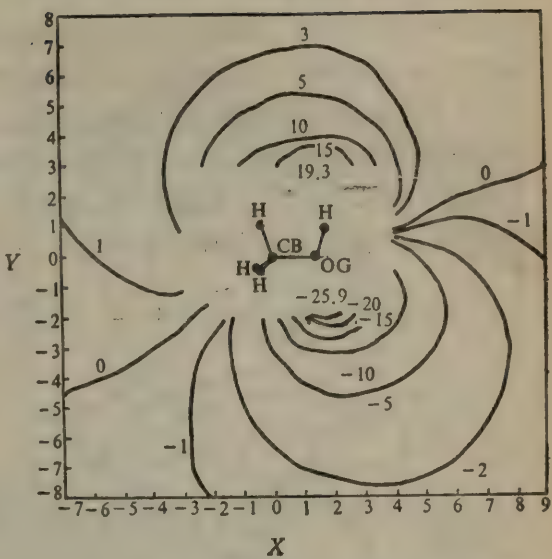


图 5-56 丝氨酸 Ser 176 侧链的分子静电势
(kcal/mol)

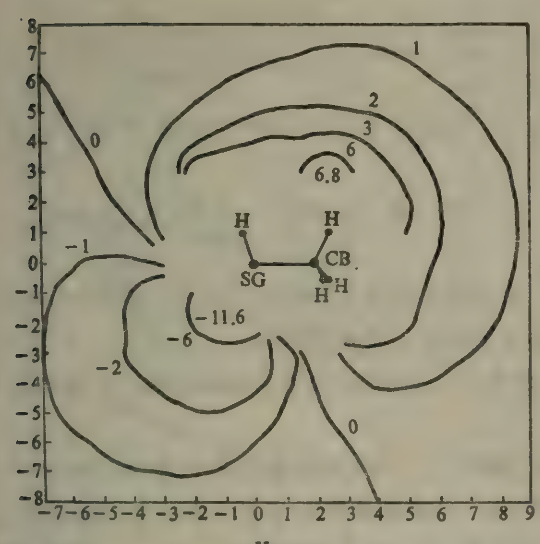


图 5-57 半胱氨酸 Cys 25 侧链的分子静电势
(kcal/mol)

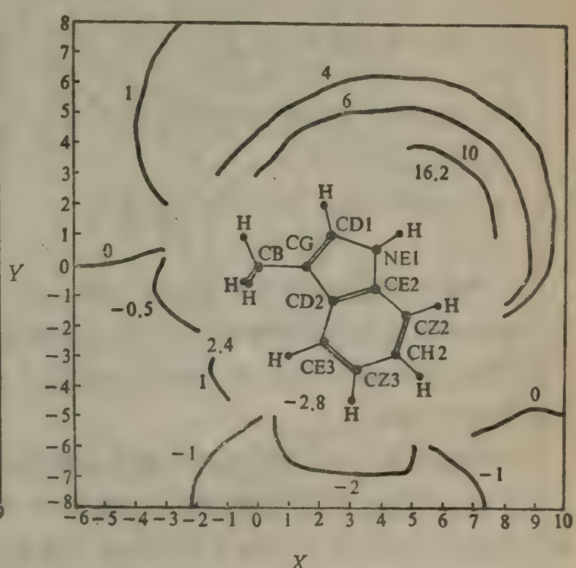


图 5-58 色氨酸 Trp 177 侧链的分子静电势
(kcal/mol)

半胱氨酸侧链在蛋白质的结构和反应中往往能起重要作用。

带电荷的侧链 带电荷的侧链包括 2 个酸性的天冬氨酸和谷氨酸，2 个碱性的赖氨酸和组氨酸。这些带电荷的酸性或碱性侧链，在溶液中既可以以带电荷的形式存在，也可以以不带电荷的形式存在，这取决于溶液环境的 pH 值。在酸性条件下，天冬氨酸、谷氨酸有不带电荷的羰基

基团，而组氨酸、赖氨酸和精氨酸则会被质子化而带正电荷。在碱性条件下，天冬氨酸和谷氨酸的羟基会电离而带负电.组氨酸、赖氨酸和精氨酸则保持不带电荷，这些氨基酸的 pK 值如下：天冬氨酸 3.86；谷氨酸 4.25.组氨酸 6.00，赖氨酸 10.53；精氨酸 12.48。在正常的生理条件下即 pH 等于 7 左右时，天冬氨酸和谷氨酸将电离成带负电荷的离子型形式，而赖氨酸和精氨酸则电离成带正电荷的离子形式，这时组氨酸则大部分不电离而保持中性。

作者计算了精氨酸、天冬氨酸 153、赖氨酸 174 和谷氨酸 50 带电荷形式残基或侧链的分子静电势，其结果如图 5-59，60，61 和 62。

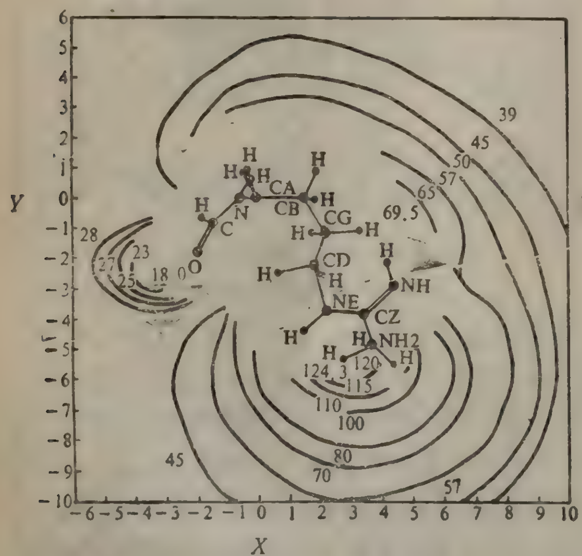


图 5-59 精氨酸 Arg 41 残基的分子静电势 (kcal/mol)

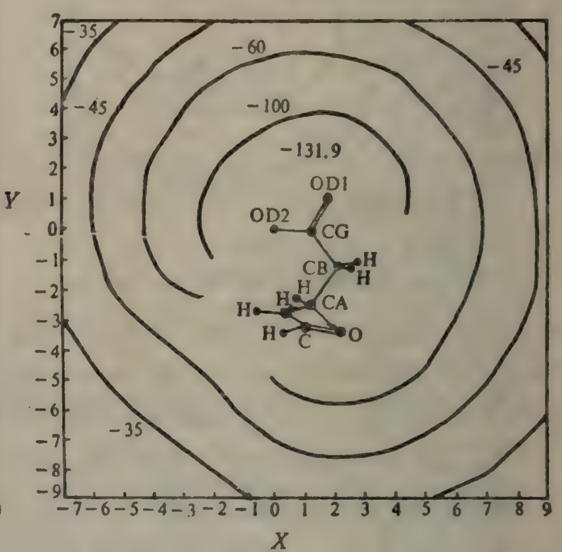


图 5-60 天冬氨酸 Asp 158 残基的分子静电势 (kcal/mol)

2 个碱性侧链由于带有羧基基团COO^-，故产生很强的负静电势场(−110——130kcal/mol)，天冬氨酸负离子比谷氨酸负离子产生更强的负静电势是由于其侧链较短。

在蛋白质的多肽链中，其溶液状态还有另两个带电荷的单元，即其头部的氨端（N 端）和尾部的羧端(C 端)。

作者还计算了木瓜蛋白酶中头部氨基酸异亮氨酸 Ile 1 和尾部氨基酸天冬酰胺 Asn 212 2 个残基的分子静电势，其结果分别列于图 5-63 和 64。氨基和羧基能分别产生一个相当强的正静电势场或负静电势场，因而它们在化学上也是活泼的，可以用某种化学反应如2、4-二硝基氟苯法和肼解法等来测定它们。

在某些蛋白质分子中，如胰岛素，其端部氨基酸的取代，尤其是其酸碱性的改变，可导致其生物活性有很大的变化，这也可从其静电势场强度变化影响了其反应性这个角度来理解。

再考察一下组氨酸的分子静电势是很有意义的，因为组氨酸在蛋白质的结构尤其是在反应中经常起着重要作用，或者是活性中心组成部分，或者作为金属离子的配位体。图 5-65 是组氨酸 159 残基的分子静电势。可以看出，在咪唑环上 NE 2 氮原子上的孤对电子产生一个很强的负

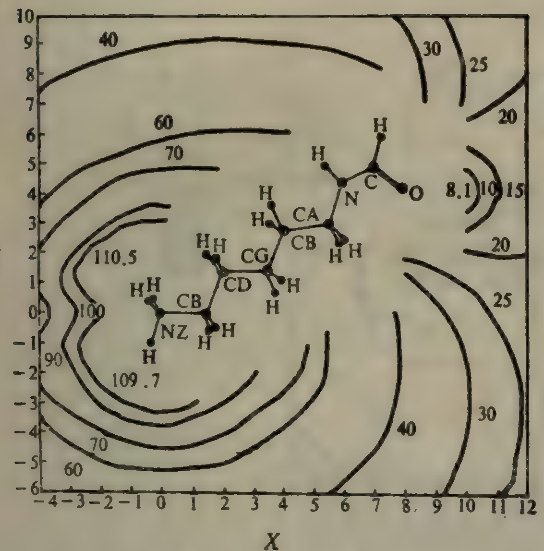


图 5-61 赖氨酸 Lys 174 残基的分子静电势 (kcal/mol)

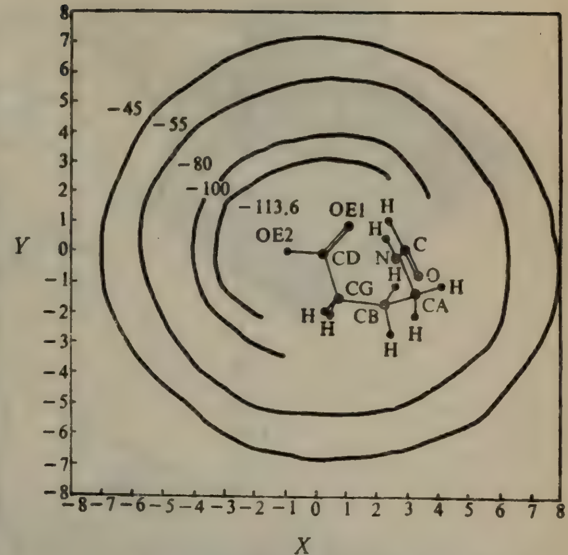


图 5-62 谷氨酸 Glu 50 残基的分子静电势 (kcal/mol)

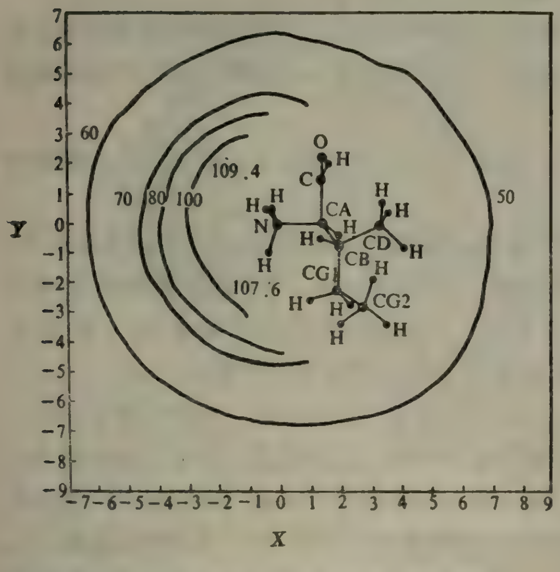


图 5-63 异亮氨酸氨端 Ile 1 残基的分子静电势 (kcal/mol)

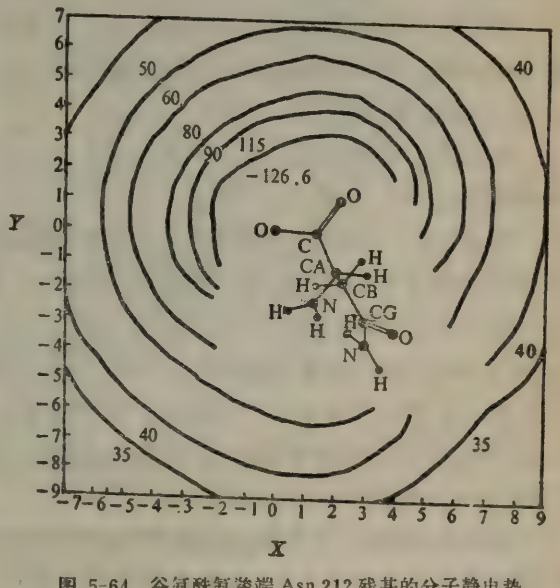


图 5-64 谷氨酰氨羧端 Asn 212 残基的分子静电势 (kcal/mol)

静电势，它甚至比肽基和侧链中的羰基氧原子所能产生的负场还要大。实际上它成为蛋白质结构中产生最强负电势的部位。此外，这对孤对电子和咪唑环上的共轭电子云使组氨酸象半胱氨酸一样成为一种强有力的金属配位体。在铁红血素、血红蛋白和细胞色素 c 等蛋白质中，NE 2 处的负势和同一个咪唑环上 ND 1 处的正静电势，常常使组氨酸成为木瓜蛋白酶、胰蛋白酶和羧

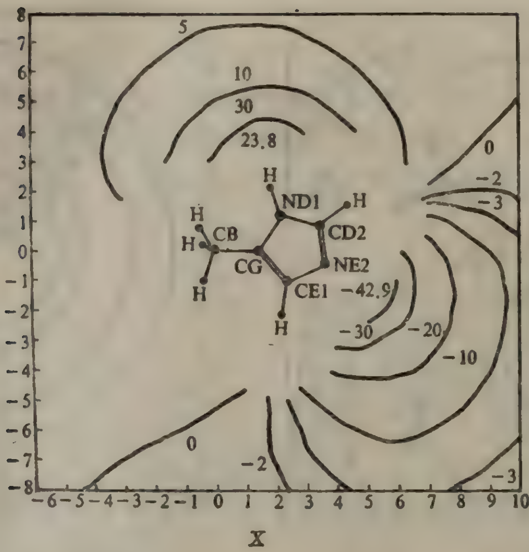


图 5-65 组氨酸 His 159 侧链的分子静电势(kcal/mol)

肽蛋白酶中"电荷转移系统"型活性中心实现质子转移的主体组成之一。

考虑到带电荷的侧链的重要性，我们对它们进行了类似肽基的构象角处理。所有的精氨酸侧链被分成二组，以精氨酸 41 和精氨酸 191 为代表(共 10 个精氨酸)，谷氨酸(共 7 个)以谷氨酸 50 和谷氨酸 89 代表。

完成了上述计算后，便可开始讨论各种次单元对整个木瓜蛋白酶分子和其活性中心的静电势分布的影响。

作者计算了木瓜蛋白酶中所有带电荷的侧链加上头、尾离解后(带电荷)的氨基酸残基在 XZ 平面方向朝 Y 面分子表面外壳上所产生的分子静电势，其结果列于图 5-66 中。其中最大值为 241.4 kcal/mol，最小值为 30.7 kcal/mol。当加上 212 个肽基单位时，结果表明，最大值(238.2 kcal/mol)和最小值(28.2 kcal/mol)和其相对位置几乎都没有什么变化。

为了了解在活性中心周围哪些氨基酸最靠近以及它们可能产生的影响，我们计算了活性中心点(活性中心假设为半胱氨酸 25、组氨酸 159 和天冬酰胺 175 这三个残基所在原子坐标的几何平均，该点坐标为 $x=19.002, y=30.624, z=17.698$)和所有氨基酸 CA 原子间的距离，离活性中心点最近邻的 24 个中性氨基酸及它们和活性中心点的距离列于表 5-16。我们可以发现，在距离活性中心半径为 0.8 nm 的球体范围内没有一个带电荷的侧链，最近的二个带电荷侧链氨基酸赖氨酸 174 和天冬氨酸 158 分别距活性中心点 0.8545 nm 和 0.8344 nm。换言之，在以活性中心点为中心直径为 1.6 nm 的球体范围内没有一个带电荷的侧链，全为中性的氨基酸。

因此，我们选取了离活性中心点 1.0 nm 以内的 24 个中性氨基酸即 2 个甘氨酸(甘氨酸20，甘氨酸 23)，4 个丙氨酸(丙氨酸 27、丙氨酸 136、丙氨酸 160、丙氨酸 162)，3 个缬氨酸(缬氨酸132，缬氨酸133，缬氨酸161)，1 个亮氨酸(亮氨酸134)，1 个异亮氨酸(异亮氨酸177)，3 个丝氨

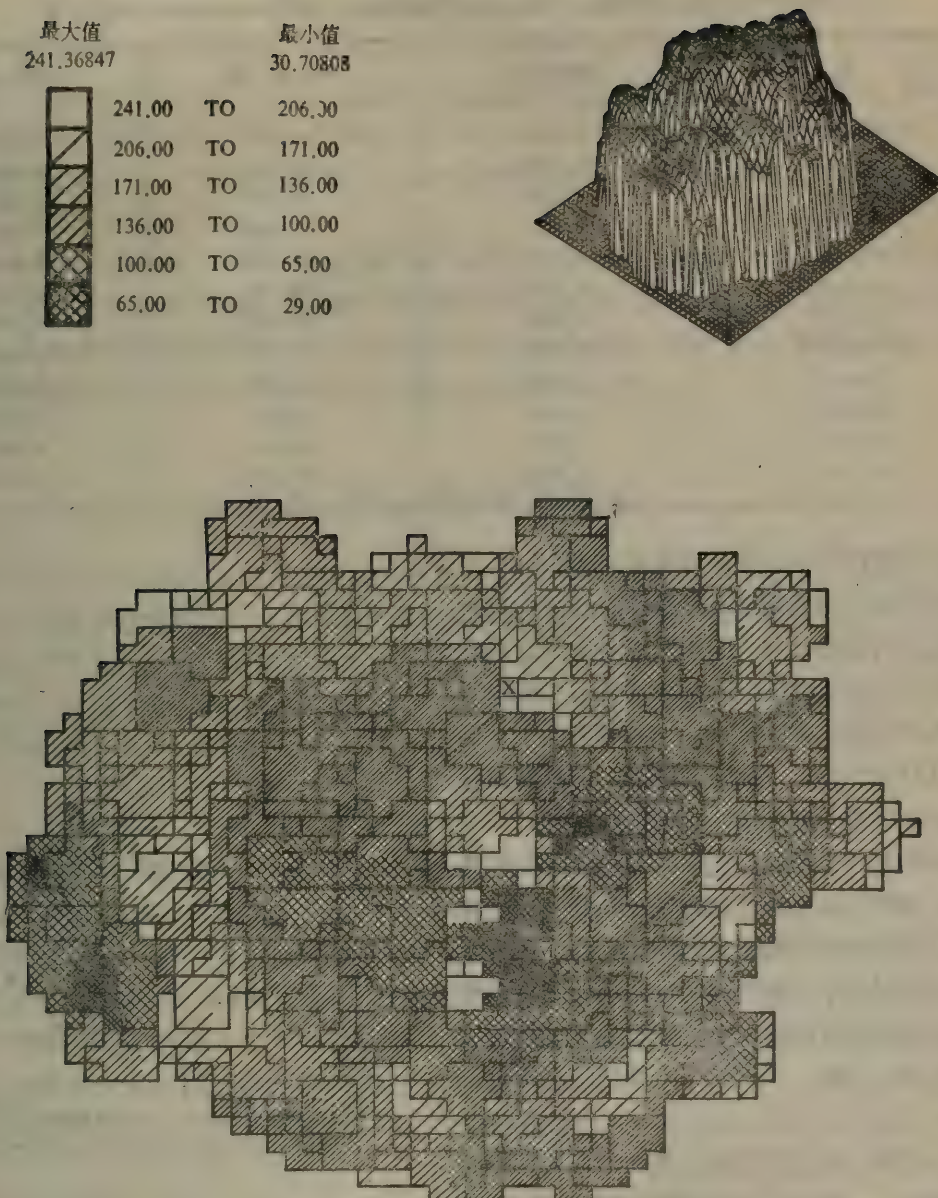


图 5-66　带电荷侧链对木瓜蛋白酶分子 XZ 平面方向朝 Y 面分子表面外壳产生的分子静电势，(kcal/mol)

酸(丝氨酸24,丝氨酸29,丝氨酸176)，2 个天冬酰胺(天冬酰胺18,天冬酰胺175)， 1 个谷氨酰胺(谷氨酰胺19)， 1 个半胱氨酸 (半胱氨酸25)，2 个苯丙氨酸 (苯丙氨酸28、苯 丙氨酸141)， 1 个组氨酸，3 个色氨酸(色氨酸26,色氨酸177,色氨酸181)，以及试验这些邻 近 氨 基 酸侧链对活性中心区域静电势分布的影响。

对活性中心和一个 XY 截面($Z=16.0$)上的静电势计算结果表明(图 5-67)：在活 性中心区

表 5-16 木瓜蛋白酶活性中心的相邻氨基酸(CA原子)与活性中心间的距离

氨基酸	距离(nm)	氨基酸	距离(nm)
His 159	0.3227	Ala 162	0.7125
Ala 160	0.3680	Phe 28	0.8937
Cys 25	0.4423	Val 132	0.8800
Ser 176	0.4452	Val 133	0.8327
Val 161	0.4481	Phe 141	0.8813
Asn 175	0.4994	Gly 23	0.8999
Trp 177	0.5481	Ile 173	0.9027
Gln 19	0.7115	Trp 181	0.9344
Leu 134	0.7345	Trp 26	0.9560
Ala 136	0.7601	Ala 27	0.9605
Ser 29	0.7606	Asn 18	0.9637
Ser 24	0.7886	Gly 20	1.0004

域有一种偶极型的分子静电势分布。图 5-67 中没有数据点的地点说明在距离某一原子0.2nm范围内。在天冬酰胺 175 周围有一个相当强的负电场，−86.4kcal/mol，该数值比一般羰基氧周围的静电势要强一倍多。它将吸引质子（氢原子)以形成氢键，而在半胱氨酸 25 周围则有一正静电势场，35.1kcal/mol，这也比任何一个孤立氨基酸中原子能产生的正静电势要大得多。它将排斥质子并帮助质子从半胱氨酸 25 的 SH 上转移到组氨酸 159 的咪唑环上。对活性中心区域分子表面外壳静电势的计算也表明了静电势最大值确是在半胱氨酸的 SH 周围。这就从理论上进一步证明了电荷转接的催化机制。

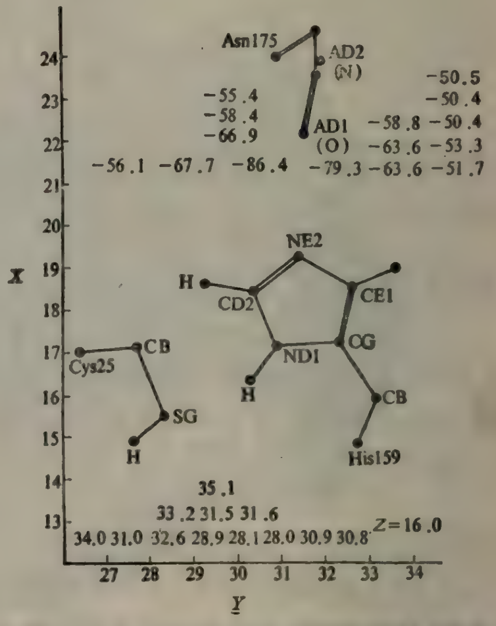


图 5-67 木瓜蛋白酶活性中心的相邻氨基酸侧链对活性中心区域产生的分子静电源（kcal/mol)

在计算过程中我们还发现在木瓜蛋白酶的晶体结构中，活性中心区域的组氨酸 159 中咪唑环的位置是 NE2(不带氢)指向天冬酰胺 175 中的 AD1 氧原子，而 ND1（带氢原子）指向半胱氨酸 25 的 SH。显然，从我们所算得的静电势分布可知，在溶液中实际发生反应时，由于天冬酰胺 175 区域对 NE2(产生负静电势)及半胱氨酸 25 处正电场对 ND1（氢产生正静电势)的双重排斥作用，很容易使组氨酸 159 或者转动单键，如 CA-CB 间的碳-碳单键，或者因咪唑环上的共轭电荷使 NE 2 和 ND1 交换氢原子而达到图 5-65、5-67 所示的能起催化作用的位置。

我们还必须注意到的是，在木瓜蛋白酶中，活性中心区连结着一段相当长的 α 螺旋，即从丝

氨酸 24 一直到异亮氨酸 40，Dujnnen 等曾非常强调这段α-螺旋对酶催化所起的作用。

如前所述，α-螺旋的作用好象一个大偶极，其 C 端带负电而N端带正电，我们计算了这段α-螺旋对整个木瓜蛋白酶分子静电势分布的影响。计算结果表明，这段长 17 个肽基的 α-螺旋对分子表面外壳层大约贡献 63 kJ/mol左右的静电势场强度。由于这段 α-螺旋是以其 N- 端连结活性中心区域半胱氨酸 25一头的正电场区域，因此它就加强了这个电荷转接系统的正电场，有利于推动 SH 上的质子转移过程。

总之，木瓜蛋白酶存在着活性中心区域，该区域的几何特征为：活性中心位于分子表面但又位于表面一条很深的沟缝的底部，底物通过嵌插作用与之反应。该区域的电荷分布特征为：质子转移是在一个电荷转接系统中进行的，而活性中心的环境，肽基链，侧链和 α-螺旋的构象造成了一个偶极型的静电势分布以利用这种质子转移而实现酶催化过程。

5.5 蛋白质的能带结构

早在 40 年代，就发现了在蛋白质内似乎存在着高效率输送能量的某种装置。例如，在肌红蛋白(Mb)的正铁血红素部分结合了一氧化碳的复合体，用对应于这个复合体的最大吸收位置的 546 nm 光照射时，若量子产额(即生成物质的分子数/被吸收的光量子数)为 1，一氧化碳就能游离。这表明，复合体所吸收的光毫无保留、毫无损失地把肌红蛋白的 Fe—CO 键切断：

$$\text{Mb—CO} \xrightarrow{h\nu} \text{Mb} + \text{CO}$$

若以 280 nm 的光照射该复合体，量子产额亦为 1，一氧化碳也能游离。由于在肌红蛋白部分中酪氨酸和色氨酸残基吸收 280 nm 的光，在正铁血红素部分理应不吸收。上述事实暗示着远离正铁血红素的蛋白质部分所吸收的光能，可能没有丝毫损失地输运到了正铁血红素部分。此外，若以 CN 基作用于具有四个正铁血红素的过氧化氢酶，恰好一个正铁血红素被 CN 基结合。如此，不仅结合 CN 基的血红素，连其他的正铁血红素也全部失去活性。这个现象称之协同性(cooperativity)或变构性(allostericity)。它也暗示着正铁血红素之间存在着某种能量传递机构。

又如，金属离子或 ATP 作用于肌肉蛋白时，仅由于 1—2 个小分子的作用，就能使蛋白质的整个行为发生变化。它也表明在蛋白质内某一位置的相互作用的能量非常有效地在蛋白质中传导。

为了说明这些实验事实，Szent-Györgyi 预言蛋白质存在能带结构。

Evans 和 Gergely 考虑到蛋白质主链间肽键往往通过氢键相互作用，存在如图 5-68 所示的以—CO—NH—为单位[肽单位]的重复结构，故有组成能带的可能性。

根据上述模型，他们用 Hückel 方法进行了计算。

由于 Hückel 方法只涉及体系的 π 电子，故在计算时，定义。单位肽中，C 有 1 个 π 电子，O 有 1 个，N 有 2 个，形成了轨道 4 π 电子系统，它们的能级图如图 5-69 *A* 所示。2 个肽单位聚集起来的能级如图 5-69 *B*。非常多的能级聚集起来的系统就形成图 5-69 *C* 那样的带结构。对应于被占轨道、空轨道各产生被占带、空带。

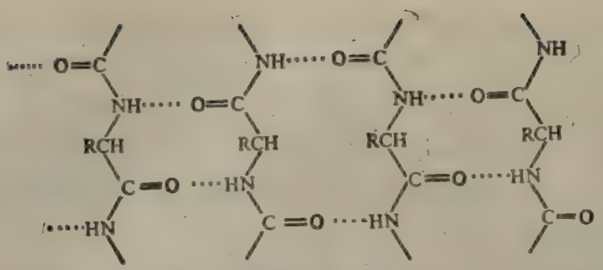


图 5-68 蛋白质主链间氢键

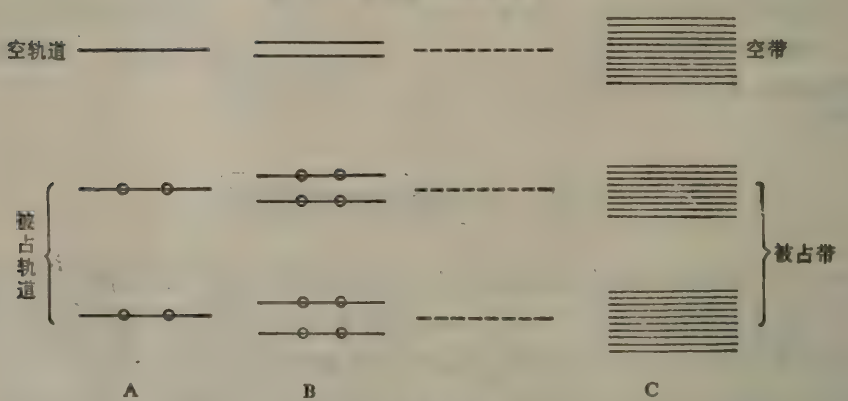


图 5-69 能级聚集成带

A. 一个单位；B. 二个单位；C. 多个单位

Evans 等计算表明最高被占带和空带的能量差约为 3 eV。Eley等关于干燥蛋白求得的能带间隔，肌红蛋白是 2.75 eV，聚甘氨酸是 3.12 eV，均在 3 eV附近。这似乎表明 Evans 的结构是正确的。但是根据 Evans 的计算，蛋白质分子在 410 nm 附近必须有吸收带，然后实验却观察不到。此外 Evans 等，包括 yomosa 所用的计算方法缺乏一般性，渐渐地被建立在固体理论基础之上的晶体轨道法(crystal orbital mothod)所代替。

下面我们用 CNDO/2 晶体轨道法研究氨基酸残基侧链对蛋白质能带结构的影响，然后讨论蛋白质构象变化对蛋白质能带结构的影响，最后讨论水对蛋白质能带结构的影响。有关金属离子的影响见 12.4.7。

由于周期的 poly(Gly)最接近蛋白质分子的骨架，故在比较氨基酸残基侧链对蛋白质能带结构影响时一般均取它为参考系。

图 5-70 是蛋白质 α-螺旋示意图；图 5-71 是聚甘氨酸能带示意图。

图 5-72 是周期蛋白质模型 poly(Ala)、poly(Asp)、poly(Gln)和 poly(His) 在与螺旋轴垂直的平面上的投影。

表 5-17 和表 5-18 分别给出了计算所得的价带、导带和禁带。

图 5-73 是所有模型系统的价带和导带的色散曲线。

从表中可以看出，与 poly(Gly)相比，除了 poly(Ala)和 poly(Asp)的价带外，所有其它系统的导带和价带都变窄了。对于导带，poly(Ala)、poly(Gln) 和 poly(His) 的带宽都只有 poly(Gly)导带宽的大约 1/2 。价带的变化则不统一：poly(Asp)和poly(Ala)的价带宽分别比

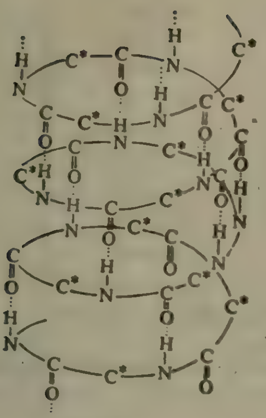


图 5-70　蛋白质的 α-螺旋



图 5-71　聚甘氨酸能带示意图

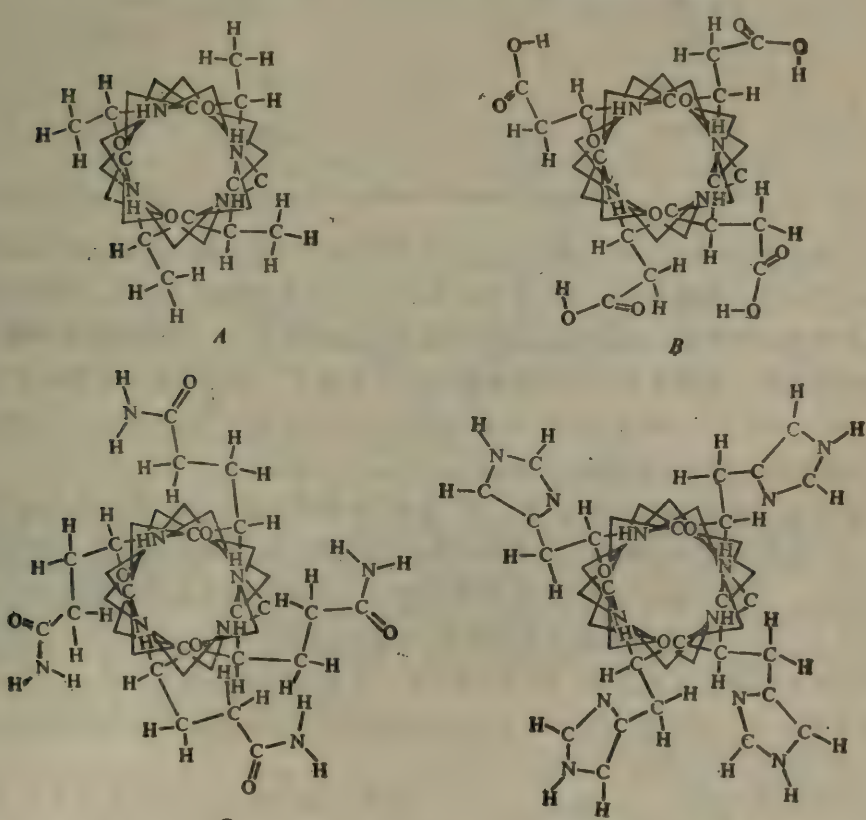


图 5-72

A. poly(Ala)、*B*. poly(Asp)、*C*. poly(Gln)和 *D*. poly(His)在与螺旋轴垂直的平面的投影图

表 5-17 poly(Gly)、poly(Ala)、poly(Asp)、poly(Gln) 和 poly(His)的导带和价带(单位：eV)

系统	带型	E_{min}	E_{max}	AE
poly(Gly)	n^*+1	−1.143	−0.653	0.490
	n^*	−11.945	−11.564	0.381
poly(Ala)	n^*+1	−1.633	−1.415	0.218
	n^*	−12.435	−11.755	0.680
poly(Asp)	n^*+1	−2.531	−2.286	0.245
	n^*	−12.925	−12.489	0.436
poly(Gln)	n^*+1	−2.122	−1.877	0.245
	n^*	−11.945	−11.673	0.272
poly(His)	n^*+1	−1.497	−1.273	0.218
	n^*	−10.830	−10.694	0.136

n^*和n^*+1分别表示价带和导带；和E_{max}、和E_{min}分别为能带的上边和下边的能量值，AB为带宽。

表 5-18 poly(Gly)、poly(Ala)、poly(Asp)、poly(Glu)和poly(His)的禁带宽度(单位：eV)

系统	禁带宽度
poly(Gly)	10.421
poly(Ala)	10.122
poly(Asp)	9.958
poly(Gln)	9.551
poly(His)	9.197

poly(Gly) 增加了0.055 eV 和0.299 eV， 后者约增加了78%；poly(Gln)和 poly(His) 的价带宽则比 poly(Gly)分别减小了0.109 eV 和0.245 eV， 后者约减小了64%。除poly(His)外，其它系统的价带都比导带宽，poly(Ala)的导带宽只有价带的 1/3 。但我们还可以看到，随着侧链结构复杂性的增加，导带宽从小于价带宽变到大于价带宽。这可以从导带和价带宽的比值看出：poly(Ala)～0.32，poly(Asp)～0.56，poly(Gln)～0.90，poly(His)～1.60。但最简单的poly(Gly)是例外，这个比值约为 1.29 。

关于价导和导带之间的禁带宽度，所有系统都比 poly(Gly)窄，减小量从约0.3 eV 到1.2 eV。而且禁带宽度减小的程度和侧链结构的复杂性有很好的对应关系。侧链结构越复杂，禁带宽度越小。例如与 poly(Gly)的禁带宽度相比，具有最简单侧链结构的 poly(Ala)的禁带宽度只减小了～0.3 eV，而具有最复杂侧链的 poly(His)则减小了～1.2 eV。一般说来，CNDO近似的缺点是计算的激发能过大，Beveridge 等人曾用 MINDO 近似计算 poly(Gly) 的禁带宽度，得到其值为 5.7 eV。若 MINDO 近似计算的能带结构也按这样的方式变化，禁带宽度将减小到 4 eV 左右，这已接近半导体的范围了。

根据 Brillouin 提出的蛋白质理论模型，氨基酸残基的侧链对主链起着杂质的作用。侧链可以作为电子受体或供体，主链和侧链间的电荷转移使主链产生空穴或多余电子，从而增加主链

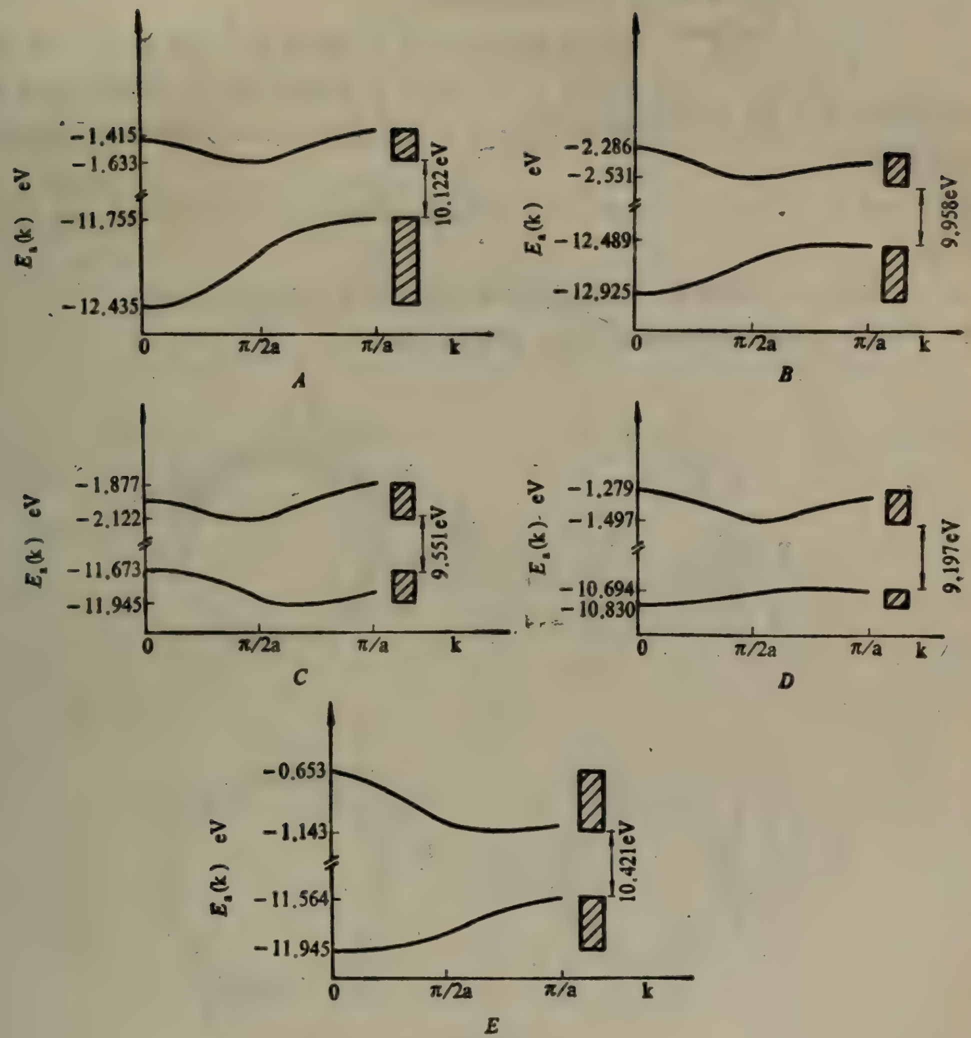


图 5-73

A. poly(Ala)、*B*. polp(Asp)、*C*. poly(Gln)、*D*. poly(His)和*E*. poly(Gly)价带和导带的色散曲线

的电导性。根据 Mulliken 集居数分析，可以计算出主链亚单位(由肽平面上的四个原子和CA-HA 组成)与侧链间的电荷转移。与 poly(Gly) 相比，poly(Ala)、poly(Asp) 和 poly(Gln)、poly(His) 的氨基酸残基的侧链都是作为电子受体，大约有0.17个电子从主链转 移到侧链。这里电荷转移的量虽然是个分数，但几个侧链的总和就能够转移一整个电子，整个蛋白质就可以产生许多带正电荷空穴，从而增加其电导性。

总之，侧链结构越复杂，蛋白质分子的禁带宽度越小。这样复杂侧链结构的蛋白质分子比具有较简单侧链结构的蛋白质分子更接近半导体的性质。另一方面，由于主链和侧链间的电荷转移，将会增加蛋白质主链的电导性。但是，侧链复杂性增加使带宽变窄，使电子(或空穴)的

有效质量变大，不易迁移。

下面讨论蛋白质构象变化对其能带结构的影响。

蛋白质构象变化是丰富的，例如 α-螺旋的螺距往往容易改变，可形成 3.0_{10}-螺旋、$3.6_{12}(\alpha)$-螺旋、3.8-螺旋、$4.4_{16}(\pi)$-螺旋和 5.2_{19}-螺旋等 α-螺旋系列结构。它们可以看成为 3.6_{12}-螺旋被拉伸或压缩后对应于不同螺距的结构。除 3.8-螺旋外，其它四种螺旋体的氨基酸残基间都有氢键形成。

图 5-74 是 3.0_{10}-、3.6_{12}-、3.8-、4.4_{16}-和 5.2_{19}-螺旋 poly(Gly) 在垂直于螺旋轴平面上的投影图。

表 5-19 给出了它们的螺距、相邻两氨基酸残基的距离和每晶胞总能量。

表 5-20 给出了它们价带和导带位置、带宽的比较。

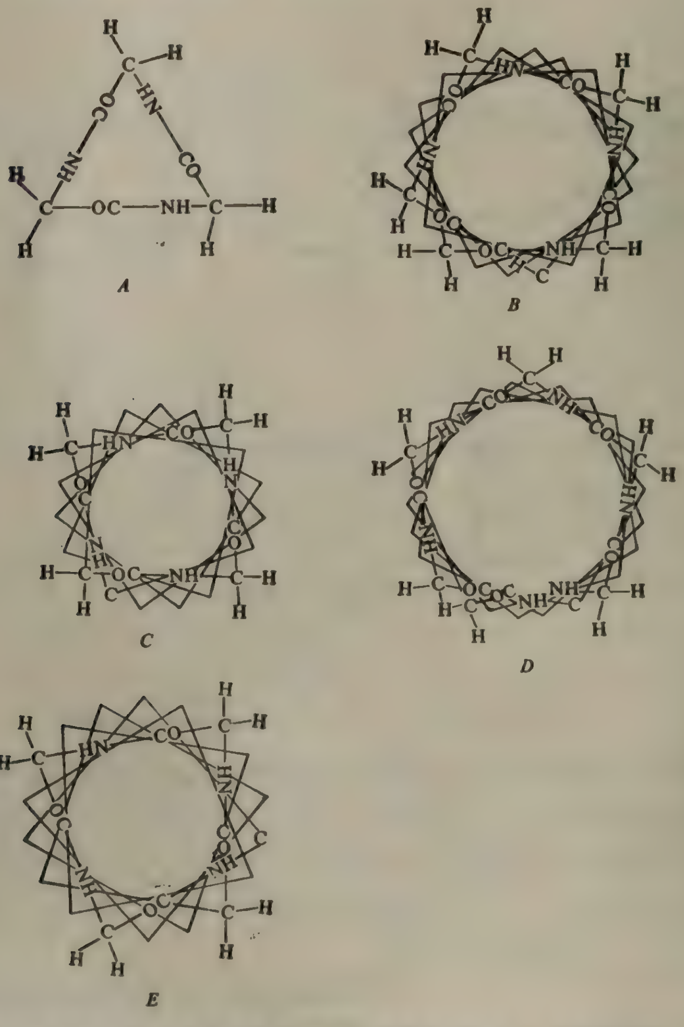


图 5-74 *A*. 3.0_{10}-螺旋体，*B*. 3.8-螺旋体，*C*. 4.4_{16}-螺旋体，*D*. 5.2_{19}-螺旋体和 *E*. 3.6_{12}-螺旋体 poly(Gly)在与螺旋轴垂直的平面的投影图。

表 5-19 3.0_{10}-,3.6_{12}-,3.8-,4.4_{16}-,5.2_{19}-,螺旋体poly(Gly)的螺距、相邻晶胞间距、每晶胞总能量

系　　统	螺距(nm)	相邻残基间距(nm)	单位晶胞总能量(au)
3.0_{10}-螺旋体	0.600	0.200	−46.539
3.6_{12}-螺旋体	0.540	0.150	−46.222
3.8-螺旋体	0.532	0.140	−46.177
4.4_{16}-螺旋体	0.506	0.115	−46.234
5.2_{19}-螺旋体	0.594	0.095	−46.284

表 5-20 3.0_{10}-,3.6_{12}-,3.8-,4.4_{16}-,和5.2_{19}-螺旋构象poly(Gly)的导带和价带(单位：eV)

系　　统	带　　型	E_{min}	E_{max}	ΔE
3.0_{10}-螺旋体	n^*+1	2.503	3.428	0.925
	n^*	−13.278	−11.945	1.333
3.6_{12}-螺旋体	n^*+1	−1.524	−1.143	0.381
	n^*	−11.755	−11.319	0.436
3.8-螺旋体	n^*+1	−2.503	−1.660	0.843
	n^*	−10.966	−10.421	0.545
4.4_{16}-螺旋体	n^*+1	−0.707	−0.544	0.136
	n^*	−12.299	−10.530	1.769
5.2_{19}-螺旋体	n^*+1	−0.490	0.327	0.817
	n^*1	−12.680	−10.367	2.313

注：n^*和n^*+1分别表示价带和导带。

表 5-21 3.0_{10}-,3.6_{12}-,3.8-,4.4_{16}-,和5.2_{19}-螺旋构象poly(Gly)的禁带宽度(单位：eV)

系　　统	禁带宽度
3.0_{10}-螺旋体	14.448
3.6_{12}-螺旋体	9.795
3.8-螺旋体	7.918
4.4_{16}-螺旋体	9.823
5.2_{19}-螺旋体	9.877

表 5-21 是它们的禁带比较。

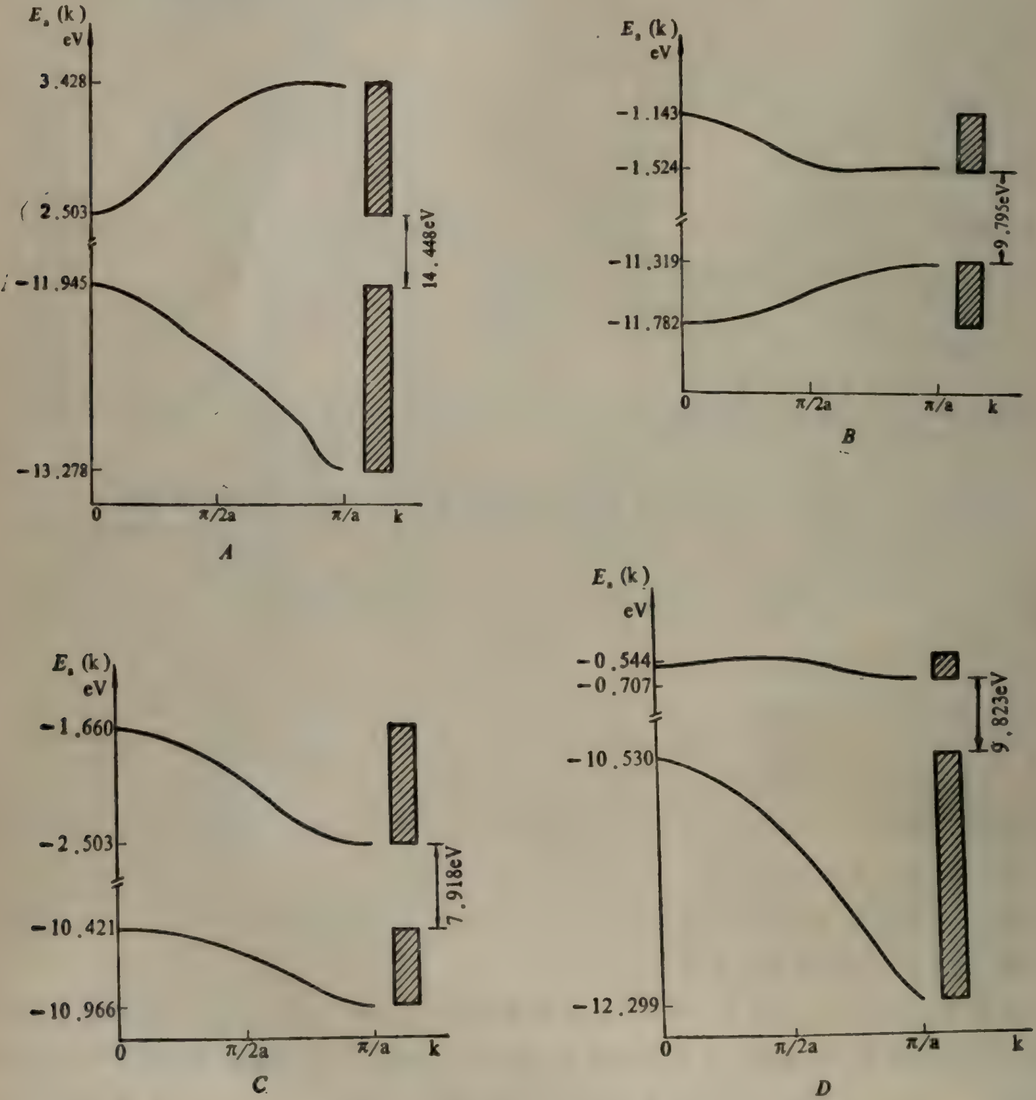
图 5-75 是价导和导带的色散曲线。

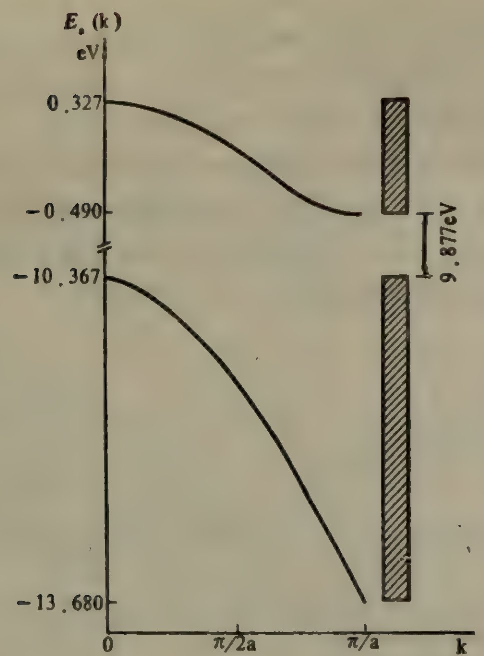
图 5-76 是它们的原子净电荷。

从能带结构表中可以看到，与 3.6_{12}-螺旋体相比，其它螺旋体的占有带平均来说都变宽了，空带则无明显改变。这里最引人注意的是价带的变化。对于 5.2_{19}-螺旋，价带宽比 3.6_{12}-螺旋体增加了 1.877eV。其它几种价带也有相当显著的增宽，3.0_{10}-、3.8-和4.4_{19}-螺旋体比3.6_{12}-

螺旋体分别增加了～0.897 eV、～0.109 eV和～0.1333 eV。导带变化比较小，约改变0.218—0.544 eV。除3.8-螺旋体外，其它四种螺旋体的价带都比导带要宽。从3.0_{10}-到5.2_{19}-螺旋体，导带和价带之比分别约为0.69、0.87、1.55、0.35。因此，不管螺距变大还是变小，能带都有显著的改变，特别是价带。改变后的价带宽要比3.6_{12}-螺旋体的约大一个数量级。因此在价带产生空穴时，其有效质量也要减小一个数量，这就更易迁移。这里给出的结果也说明蛋白质与电子受体相互作用，在价带内产生空穴，能使能量和电荷长距离迁移。

从表 5-21 可以看出，当螺距增大时(相对3.6_{12}-螺旋体而言)，禁带宽度变大。当螺距减小不多时，即3.8-螺旋体时，禁带宽度约减小1.877 eV。这个变化是相当大的。但当继续减小时，禁带宽度反而增加，不过与 3.6_{12}-螺旋体很接近，仅相差～0.05 eV。





E

图 5-75 *A*. 3.0_{10}-螺旋、*B*. 3.6_{12}-螺旋、*C*. 3.8-螺旋、*D*. 4.4_{16}-螺旋和*E*. 5.2_{18}-螺旋 poly (Gly) 的价带和导带的色散曲线

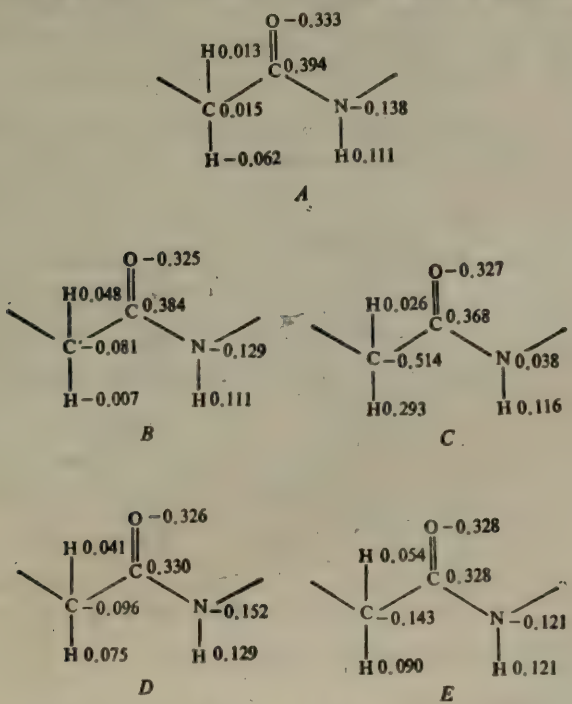


图 5-76 *A*. 3.0_{10}-螺旋，*B*. 36_{12}-螺旋，*C*. 3.8-螺旋，*D*. 4.4_{16}-螺旋和*E*. 5.2_{18}-螺旋 poly(Gly)晶胞中的原子净电荷(单位:au)

在五种螺旋体中，原子净电荷都基本相同。这表明它与螺距的改变没有多大关系。

综上所述，一方面螺距变化不大时，带宽变化不大，但禁带宽度大大减小，有可能使蛋白质趋向半导体，另一方面，虽然螺距增大或减小很多时，禁带宽度都是增大，但价带的显著变宽使得蛋白质与电子受体作用时，产生的空穴的迁移率可能很大。

最后，我们讨论水对蛋白质能带结构的影响。我们只研究水与肽平面上的原子的相互作用，而不考虑与侧链原子的相互作用，并且只研究水结合到肽平面上不同原子对蛋白质能带结构的影响。计算针对两种水和 poly(Gly)作用的复合体。每个晶胞中包括一个甘氨酸残基和一个水分子。水分子与甘氨酸残基的相对位置如图 5-77、5-78 所示。

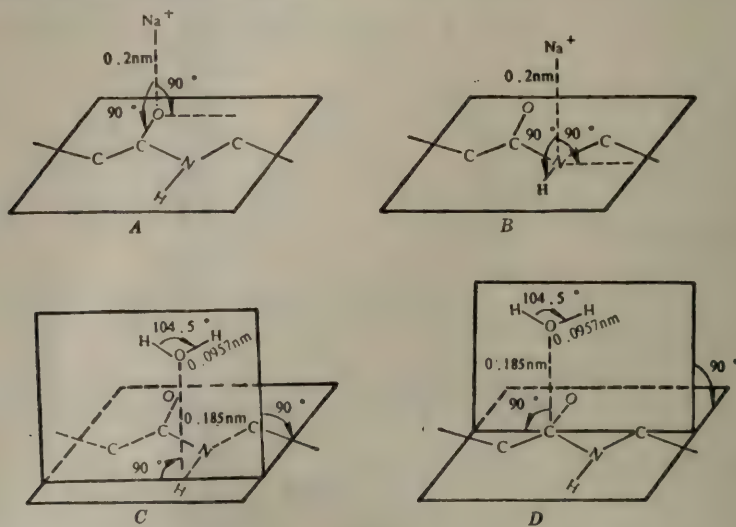


图 5-77 *A*. poly(Gly-Na^+)I，*B*. poly(Gly-Na^+)II，*C*. poly(Gly-H_2O)I，*D*. poly(Gly-H_2O)II中的离子或水分子与肽平面的几何关系

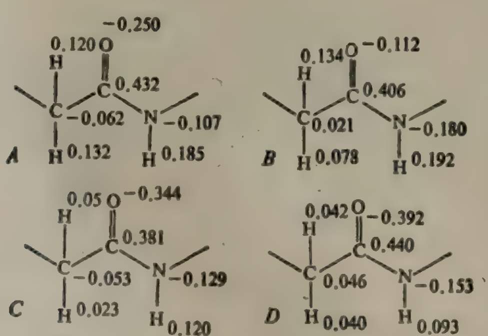


图 5-78 *A*. poly(Gly-Na^+)I，*B*. poly(Gly-Na^+)II，*C*. poly(Gly-H_2O)I，*D*. poly(Gly-H_2O)II的原子净电荷(单位：a.u.)

表 5-22 是它们的价带和导带的位置、带宽的比较。

表 5-23 是它们禁带宽度的比较。

从表中可以看到，水与p oly(Gly)作用时，与 Na^+ 的作用相比，对能带 的位置影响不大。

表 5-22 poly(Gly-Na$^+$)I, poly(Gly-Na$^+$)II, poly(Gly-H_2O)I, poly(Gly-H_2O)II的价带和导带(分别用n^*和n^*+1表示)(单位：eV)

系统	带型	E_{min}	E_{max}	ΔE
poly(Gly-Na$^+$)I	n^*+1	−14.884	−14.775	0.109
	n^*	−26.013	−25.305	0.708
poly(Gly-Na$^+$)II	n^*+1	−16.516	−16.054	0.462
	n^*	−26.747	−26.040	0.707
poly(Gly-H_2O)I	n^*+1	−0.735	−0.272	0.463
	n^*	−11.428	−11.129	0.299
poly(Gly-H_2O)II	n^*+1	−0.544	0.000	0.544
	n^*	−11.238	−10.612	0.626

表 5-23 poly(Gly-Na$^+$)I, poly(Gly-Na$^+$)II, poly(Gly-H_2O)I,poly(Gly-H_2O)II的禁带宽度(单位：eV)

系统	禁带宽度	系统	禁带宽度
poly(Gly-Na$^+$)I	10.421	poly(Gly-H_2O)I	10.394
poly(Gly-Na$^+$)II	9.524	poly(Gly-H_2O)II	10.068

和 Na$^+$ 离子一样，能带结构的变化程度是和水结合部位有关的。当水结合到 NH 原子上时，与 poly(Gly)相比，能带和禁带的宽度改变都很小：价带宽减小了21%，导带宽增加了 5.5 %，禁带宽仅减小了0.027 eV。当水分子与肽平面上羰基的碳原子结合时，改变要大。与 poly(Gly)相比，价带宽约增加了一倍，导带宽约增加了10%，禁带宽则减小了0.353 eV。且导带比价带窄。

此外，poly(Gly-H_2O)I 和poly(Gly-H_2O)II 的羰基比 poly(Gly)极性要强。这有利于氢键形成。在这两种模型中，主链和水分子之间的电荷转移都很小，只有约0.015电子电荷。

综上所述，水与 poly(Gly)作用在不同程度上使禁带变窄，使蛋白质向半导体的范围变化。同时。水分子与 poly(Gly)作用大都使价带增宽，且价带比导带宽。这就表明蛋白质分子中的空穴有较大的迁移率。这有利于能量和电荷在蛋白质分子中的迁移。

(参考文献见第 6 章。)

邵 俊

第6章　酶作用机理的量子生物学研究

酶作为一种蛋白质,具有一般蛋白质的物理化学性质,特别是它具有一级结构、二级结构、三级结构和四级结构。然而,酶又是一类具有生物催化功能的特殊的蛋白质。作为催化剂,酶有下面两个显著特点:1. 具有很高的催化效率。在相同的条件下,酶的存在可以使一个反应的反应速度大大加快。拿一个最简单的二氧化碳水合反应来讲,$H_2O+CO_2 \rightleftharpoons H_2CO_3$,在生物体内该反应是由碳酸酐酶催化进行的。如果没有碳酸酐酶存在,二氧化碳从生物体组织到血液然后再到肺泡的空气中的反应,就会很不完全。碳酸酐酶的催化反应,是目前已知的最快的酶的催化反应之一。每个酶分子,在一秒钟内可以使十万个二氧化碳分子发生水合反应,酶的催化反应比非酶反应要快千万倍。尿素被脲酶催化水解反应比非酶反应要快 10^{14}倍。表 6-1 是另外的一些例子。2. 具有高度的专一性。一般地讲,酶具有两方面的专一性:对于被作用的反应物(又称底物)的专一性和对于被催化反应的专一性。不同的酶,其专一性程度颇不相同。

表 6-1　酶反应和非酶反应的速度比

酶	$v_{酶}/v_{非酶}$
溶菌酶	2×10^{8}
糜蛋白酶	4×10^{3}
β-淀粉酶	3×10^{11}
反丁烯二酸酶	2×10^{11}

$v_{酶}$、$v_{非酶}$ 分别是酶反应、非酶反应的速度,后者指用酸、碱催化时的值。

几十年来,伴随着 X 衍射等各种谱仪在酶学中的应用,人们已经积累了丰富的资料,其中最重要的结果是对酶的活性部位有了更深入细致的了解。在蛋白质数据库中,已贮存有几个极其复杂的酶蛋白晶体结构数据,从中我们可以得到活性部位的几何构型和其它精细结构信息。

另一方面,量子化学已发展成为理解化学反应机理的有力工具,用半经验的或极小基的从头算法对 20—30 个第二周期元素原子组成的分子进行量子化学计算已是很方便的事情了。在研究基元过程时还发展了考虑溶剂等环境效应在内的种种方法。作为分子轨道法的一种补充,分子力学亦获得了成功。

实验和理论的进步使得深入了解酶过程成为可能。酶的量子生物学研究的目的就是要从电子水平上来说明上述问题。

6.1　模型和方法

化学现象的理论研究需要适当的模型,而模型又必须建立在大量的实验事实的基础之上。酶反应是一个极其复杂的过程,因此构造一个合适的模型所需要的实验资料要比构造一个化学

反应的模型多得多。下面，我们从方法论的观点来讨论这个问题。

6.1.1 酶反应模型

迄今，酶反应通常都是用过渡态理论来解释的。这类反应沿着反应途径的标准 Gibbs 自由能ΔG^0的典型变化如图 6-1 所示。酶反应的一个重要特征是在反应的第一阶段生成酶-底物配合物。这个配合物的稳定性是由于酶和底物间的非共价相互作用(氢键、范德华力等)所造成的。酶-底物配合物的生成已由实验证实,但在绝大多数情形，最终反应产物并不是经过一个简单的基元反应步骤由酶-底物配合物生成的。例如,丝氨酸酶对肽键的裂解要经过几个相对稳定的中间产物,这些中间产物将加速整个酶反应过程。动力学实验、X 衍射和其它结构研究都证实了这些中间物的存在。

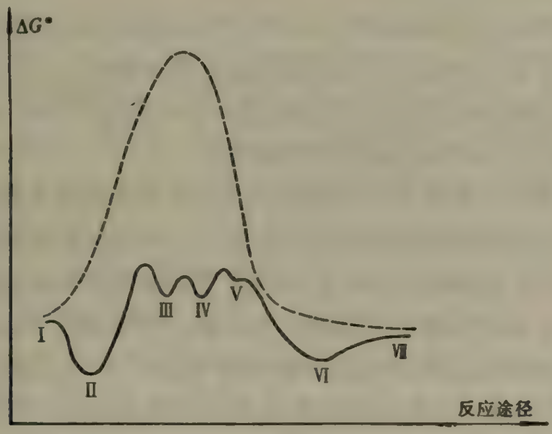


图 6-1 沿着反应途径标准 Gibbs 自由能变化实线为丝氨酸蛋白酶的反应途径，虚线表示气相中相似反应

酶蛋白一般由 200 个以上的氨基酸残基构成。也就是说，每个酶蛋白由约3 500 个原子和约 15 000 个电子组成。并已知道,酶的催化作用仅发生在酶分子上的一小部分区域，且只有极少数原子直接参与催化过程。这个一小部分区域就叫做活性部位。活性部位是有三维结构的。没有酶的二级、三级结构就不可能构成有效的活性部位,因为活性部位起作用的基团并不是由一级结构相近邻的氨基酸提供,如 α-糜蛋白酶,丝氨酸-189,丝氨酸-190，半胱氨酸-191,甲硫氨酸-192,甘氨酸-193,天冬氨酸-194,缬氨酸-213,丝氨酸-214,色氨酸-215，甘氨酸-216，丝氨酸-217,丝氨酸-218,苏氨酸-219,半胱氨酸-220,它们在一级结构中并不紧挨，然而由于多肽链的折叠,它们在三维空间中靠得很紧,并作为一个整体起作用,构成了 α-糜蛋白酶的活性部位。活性部位处在酶分子表面的一个裂罅内，在那里存在着可以和底物结合、使底物起催化反应的基团。结合基团的总和叫结合部位,催化基团的总和叫催化部位。在 α-糜蛋白酶中，丝氨酸-195的羟基、组氨酸-57 的咪唑基和天冬氨酸-102 的羧基形成了催化部位,称为催化三价基(catalytic triad),它是活性部位的中心机构。而其它侧链在酶-底物配合物的稳定和其它中间态的稳定中起了很大的作用。活性部位是高度水化的,在其周围可以找到离子对。应该指出,活性部位、

结合部位、催化部位是为讨论酶的作用机制的方便才把它们抽象出来的。实际上，它们是和整个酶分子不可分的。酶的其它部位对催化作用和调节作用、免疫作用、维持一定空间结构等方面都有很重要的作用。

研究实践表明，要用理论计算来定量描述酶的催化过程，恰当设计活性部位模型是第一重要的。为此，必须要有 X 衍射结构数据。另外，必须将催化部位从活性部位中抽提出来。最流行的方法是把活性部位分成两部分，第一部分为直接参与催化过程的中心机构，它们参与键的生成和断裂。例如，在 α-糜蛋白酶中，这些原子就是前面提到过的催化三价基。在量子化学从头计算中为使计算更为方便，又用乙醇或甲醇代替丝氨酸，用咪唑代替组氨酸，用乙酸或甲酸代替天冬氨酸。活性部位或酶的其它部分可考虑为环境。环境可用分子力学或简单的量子化学方法来处理。早期的工作只独立考察中心机构的作用，而对在催化过程中对酶与底物结合的位置和取向起作用的环境的影响仅通过间接方式引入。

水化可以通过结构模型或连续模型来模拟，离子对则可用点电荷模型。水化效应和离子对效应模拟的关键是活性部位水分子和离子对的个数及位置的确定。由于实验对水化只能提供有限的信息，必须从理论角度来构造水化层。最简单的是仅应用化学知识，即带有孤对电子的原子(N、O和 S)和水分子中的 O—H 键结合，酸性的$X-H$($X=$N、O、S 和 C)和氧的孤对电子相互作用。虽然键合作用强烈的离子对的位置用 X衍射可以定出，但酶在发挥催化功能时，这些离子对是否仍然在生物系统中、它们的变化情况等还没有最后从实验上得到证实。

为了查明酶环境的作用，通常要把酶反应和溶液中相似的化学反应比较。但对酶催化部位的中心机构进行量子化学计算所用的模型是气相中的情形(即气相模型)，因此简单对比会引出错误的结论。为了避免这类错误，必须同时对酶反应和溶液中化学反应的环境影响都进行模拟计算。在计算时，要保持模型的一致性。也就是说，若酶只考虑第一水化层，相应的化学反应也只要考虑第一水化层就够了。

6.1.2 催化部位的分子轨道计算

前面提及的催化部位的中心机构指的是酶骨架和侧链中这样一些原子的集合，它们在反应过程中直接参与键的生成和断裂。这个原子集合可以用象水、甲酸和咪唑这样一类极其简单的分子作为模型。即使如此，这样的模型至少也得包含 15～20 个原子或 60～80 个电子。由于计算机的发展，已有不少人用从头计算法来处理它们了。表 6-2 总结了这方面的一些工作。

前面已经说过，酶催化过程一般用过渡态理论来描述，因此必须对反应体系的位能面有所了解。也就是说，至少要对反应体系的几何构型进行优化。为控制工作量，对不同的对象所需优化的几何构型参数是不同的，优化参数需精心选择。

特别值得一提的是，模型和方法必须匹配。例如，若不考虑环境效应，那么即使对催化部位中心机构的气相模型作最精确的从头算，对酶的催化过程也是得不出什么深刻结论的。另外，还必须考虑从某种分子轨道理论所算得的物理量是否精确到能取代或解释实验事实。理论描述酶反应最常用并且也是最重要的性质是：1. 过渡态(位能面中的局部极小值)之间的能量差；2. 决定基元反应步骤速率的活化能；3. 为考察环境效应和反应性所需的电荷分布情况；4. 分子构

表 6-2 用分子轨道理论研究酶的例子

方法	基组	蛋白质	模型
ab initio	双 ξ	木瓜蛋白酶	MeSH···Im···HCHO[11]
ab initio	4-31G	丝氨酸蛋白酶	MeOH···Im···HCOOH[19]
ab initio	STO-4G	羧基肽酶对Zn取STO-4G+3d+4p	$Zn(OH)(NH_3)_2^+$[20]
ab initio	STO-3G	乙二醛酶I	MeCOCH(OH)SMe[22]
PRDDO	极小基	丝氨酸蛋白酶	H_2NCHO···MeOH···Im···HCOOH[24]
MINDO/3	极小基	细胞色素 P 450	CH_4或C_2H_4+O[28]
CNDO/2	极小基	α-糜蛋白酶	MeOH···S···Me···Im···MeCOOH[30]

象。

反应中间态和产物之间的能量差的计算是计算量子化学的中心问题。若原子环境和电子对的数目在反应过程中是守恒的，例如质子化，相关能效应是很小的。双 ξ 加极化基组所产生的 Hartree-Fock 极限为 5～10kJ/mol，而 4-31G 基组的误差在 ±80kJ/mol。例如，甲酸阴离子的实验质子化能是 −1430kJ/mol，用 4-31G 得到的值为 −1488kJ/mol。遗憾的是在酶反应中，对质子转移反应，要想获得如上所述的好结果是很困难的，除非用双 ξ 加极化基。然而酶催化中心机构模型含有 50～60 个电子，要做如此复杂的计算并不是件容易的事情。如果用 4-31G基的话，那么由于电荷分布的计算不够正确，所获得的相互作用能就不可靠。为了得到正确的质子转移能，必须根据实验数据对计算值作适当的修正。例如，对丝氨酸类蛋白酶中心机构中的质子化咪唑基-甲酸阴离子部分，就是采用 4-31G 基来计算其质子化能而用 STO-3G 极小基结合基于实验数据的适当修正来估计咪唑和甲酸间的相互作用能的。

若在反应过程中，电子对的数目和环境发生了变化，例如：$HCN+3H_2 \rightarrow CH_4+NH_3$或$H_2O \rightarrow OH^-+H^+$，相关能就极其重要了。由于催化部位的中心机构还是比较大的，目前要对它们做相关能计算还是不现实的。在这种情形下，可做MINDO系列的计算(MINDO/3，MNDO，MNDOC)。尽管这些方法的近似性和适应性尚在讨论之中，它们毕竟已在几个酶的研究中获得了成功。

极小基的从头算和各种半经验量子化学方法(CNDO/2、EHMO等)有它们的缺点，但是只要我们把计算结果跟恰当的实验联系起来，还是能从理论和实验的对比中得到一些基本的靠得住的信息的。最典型的例子之一是 Kitayama 和 Fukutome 对丝氨酸蛋白酶中组氨酸碱性的讨论。

活化能定量的计算需要考虑相关效应和用大基组。这只对不超过 15～20 个电子的模型是可实现的，在某些情形MINDO可获成功。由于酶反应的复杂性，基元反应步骤(系统从其势能面的一个极小点运动到另一极小点)，只有在极少数情形下才被研究过。一般，实验测得的反应速率涉及到若干个基元反应，因此要将理论活化能和实验结果比较是很困难的。从这种意义上来讲，酶反应理论研究中活化能的计算还不算最棘手的问题。

电荷分布最好从 CNDO/2 和从头算 STO-3G基获得，而 4-31G虽然基组比 STO-3G大，但

在电荷分布上的结果却不好。分子构型用 STO-3G 也能得到满意的结果，只是与氢键合的键长不够理想。CNDO/2 也能得到分子构型较好的结果，但氢键键长的计算值也不理想(偏短)。

总之，对分子不同的性质，为了得到同等程度的可靠性，需用不同的方法。在酶反应理论研究中为了得到最大的成功，在对理论和实验进行细致的比较和分析时，往往需采用各种不同的计算方法。

6.1.3 分子力学计算

分子力学的基础是该方法认为化学键性质在某些类型的分子内具有迁移性。在分子位能面局部极小点附近，势能函数总可以近似地表为一些简单的解析函数的和的形式。函数的形式及其中的参数可根据与实验数据，如平衡键长、键角、振动频率、旋转势垒等的拟合而得到。分子力学可以看作是一种内插法。只要用得适当，分子力学可以对象酶反应这样大的系统获得相当精确的结果。目前，已经有了几套用于计算酶的结构和酶-底物配合物结构的分子力学参数。

表 6-3 是将分子力学方法用于酶的研究的部分工作。

表 6-3 分子力学在酶学中的应用例子

酶	计算目的
α-糜蛋白酶	底物专一性[70]
木瓜蛋白酶	环境对催化反应的影响[7,8]
溶菌酶	抑制剂和底物与活性部位结合的优势构象[65,69,71,72]
溶菌酶	水化[73,74]
溶菌酶	反应中间物稳定性研究[10]

分子力学研究能帮助我们回答酶和底物如何相对取向、酶和底物构象如何变化其间的相互作用最大等问题。分子力学方法还可跟量子化学方法结合来研究蛋白质环境对酶催化过程的影响。

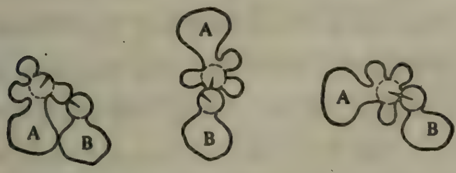
分子力学计算的最大优点是它的简单性和有效性。用该方法能对所研究对象作整体优化。势参数选择好的话，该方法可达 2kJ/mol 的精度。其缺点是势参数确定有任意性，对同一问题往往有多套参数可供用，因此用不同参数对同一系统的计算，是很难加以比较的。

6.1.4 轨道取向、轨道对称性和稳定能

继提出酶反应中邻近效应和定向效应 (orientation effort) 的重要性，Koshland 又提出了轨道取向 (orbital steering)的概念。就是说，从图 6-2 (1)的邻近效应和定向效应都没有的状态变化为具有邻近效应但无定向效应的状态，只有进一步到达具备邻近效应和定向效应时才能起反应。图 6-3 表示了图 6-2 (3)的状态中活性中心和底物反应位置的原子轨道取向模式图。Koshland 认为在酶反应中由于进行了高效的轨道取向产生了反应加速。

α-糜蛋白酶反应时，丝氨酸 195 的 OH 基解离成—O^-基“攻击”底物的羰基碳，亲核攻击阶

段是反应速度的限制步骤，因而取决于—O^-基和羰基碳位置电子轨道重迭的系统的稳定性，这可用轨道重迭的程度引起系统总能量的变化来加以考察，结果见图 6-4 。随着轨道取向变坏，能量急剧变化，显示出轨道取向效应的重要性。



(1)　　(2)　　(3)

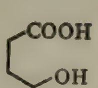
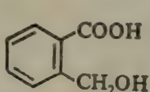
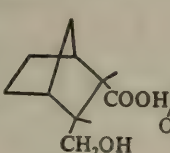
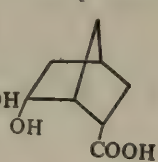
图 6-2 (1) 无邻近、定向效应；(2) 有邻近效应，无定向效应；(3) 有邻近、定向效应



图 6-3 酶活性中心与底物反应基团之间的轨道取向

表 6-4 中一系列化合物酯化的例子也说明当轨道取向越正确时反应速度越大（反应时参与反应的原子固定在一定的立体位置）。

表 6-4 各种化合物酯化速度比
（以乙酸和乙醇的酯化速度为 1）

	CH_3CH_2OH + CH_3COOH				
速度比	1	79	305	6630	1 027 000

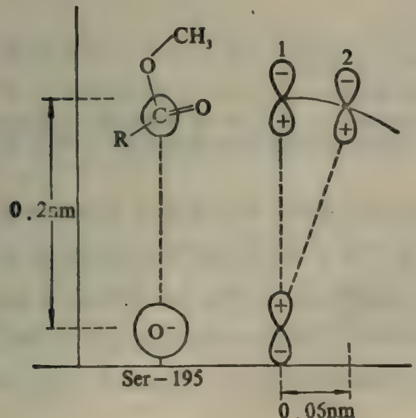


图 6-4 α-糜蛋白酶的活性中心（Ser-195）和底物羰基碳电子轨道的关系
1—轨道取向位置；2—移动 16°42′的位置

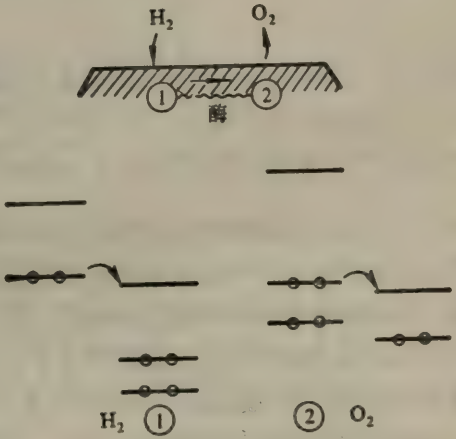


图 6-5 有关酶的电子移动模式图

Ferreira 应用轨道对称性的设想来解释酶反应机制，提出了一些颇有意义的想法。

假定酶的一个活性位置（记为 1）的最低空轨道具有 H_2 的最高被占轨道相同的对称性。紧靠旁边的另一个活性位置（记为 2）的最高被占轨道和 O_2 的最低空轨道具有相同的对称性，图 6-5。这样，电子从 H_2 的最高被占轨道向活性位置 1 的最低空轨道流动，从第 2 活性位置的最高被占轨道流向 O_2 的最低空轨道。这种电子从活性位置 1 向 2 流动的电子回路从能带来看是一个封闭回路，结果生成 $2H^+$和$(O—O)^{2-}$离子，因为是离子间的反应，从能量来讲反应容易进行。从轨道对称性来看，因为电子从$(O—O)^{2-}$的 π 轨道向 $2H^+$的 1s 轨道移动，$a_1\times a_1\rightarrow a_1$，所以是对称性容许反应。从图 6-6 中的 $H_2+O_2\rightarrow H_2O_2$ 的轨道相关图也能够很容易地解释上述反应。Ferreira 用完全相同的想法首先指出乳清酸核苷-5′—磷酸—焦磷酸化酶切断乳清酸离子的 N—H 键和 PRPP(磷酸核糖焦磷酸)的 C—O 键，是由于底物和酶的活性中心之间的电子移动从对称性禁止变成对称性容许。

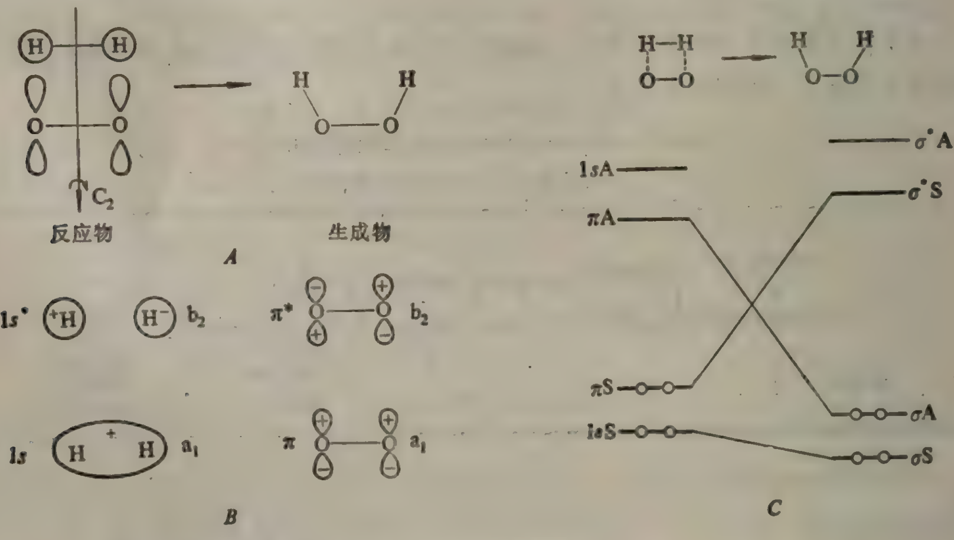


图 6-6 *A*. $H_2+O_2\longrightarrow H_2O_2$ 的模式图

B. H_2 和 O_2的成键、反键轨道的对称性

C. $H_2+O_2\longrightarrow H_2O_2$ 的轨道相关图

1s，π 分别为 H_2，O_2 的轨道，σ，σ^*是 O—Hσ 键的成键、反键轨道，

S，A 是对H—H，O—O的中心线呈镜面对称的对称性

因此，可以认为酶的作用在于通过底物和它之间的电子移动，导向对称性有利的方向，这很可能是研究酶反应的一种饶有兴趣的工作设想。只是在象 α-糜蛋白酶等以酰化化合物为中间体的酶，轨道对称性分析不能简单照搬。但是即使在这种情况下，要是考虑到轨道对称性容许，反应分子和被反应分子轨道大量重迭这一点，生成酰化中间物时酶的活性基团和底物反应位置轨道重迭也是很大的，就是说不是别的正是显示出轨道取向的重要性。在这种意义上，轨道对称性的设想和轨道取向的设想基本上是同一的东西。

稳定能是两个系统相互作用产生的能，它一般使系统趋于稳定。稳定能中起作用的力有静电相互作用，范德华力等种种，但是在酶反应中具有最重要意义的是非定域能(delocalization

energy)。

所谓非定域能是因为分子间或分子内电子移动产生的能量，所以在过渡状态附近试剂和底物之间产生电子移动时，非定域能使系统稳定，造成活化能降低。这个理论指数在有机反应机制的分析中另有极为重要的地位，在生物体反应中也趋向于广为应用。下面举几个例子简述之。

图 6-7 是胆碱酯酶水解反应的模型，表 6-5 是对过渡态附近产生酶和底物间电子移动模式的非定域能，计算与相应的实测活化焓符合得很好。

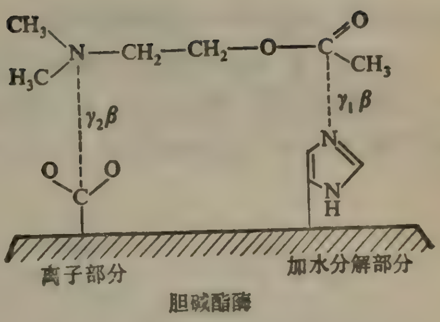


图 6-7 胆碱酯酶和底物间的电子移动

表 6-5 伴随酶-底物间 π 电子移动的非定域能 (δe)和活化焓($\Delta H^{\neq}$)的比较

底物	$\Delta H^{\neq}$ (kcal/mol)	加水分解部分 $\delta e_1(\gamma_1^2\beta)$	离子部分 $\delta e_2(\gamma_2^2\beta)$
$(H_3C)_3N^+-CH_2-CH_2-O-C(=O)-CH_3$	14～19	−0.379	0
$(H)_2N-CH_2-CH_2-O-C(=O)-CH_3$	9.5	−0.370	−0.664
$(H)(H_3C)N-CH_2-CH_2-O-C(=O)-CH_3$	8	−0.370	−0.748
$(H_3C)_2N-CH_2-CH_2-O-C(=O)-CH_3$	6.7～8	−0.370	−0.857

β为负量，所以δe为正。这个绝对值越大，活化焓越小。

溶菌酶表面上有细长的沟(称为 A,B,C,D,E,F)，这里是与底物结合的地方。键的切断有专一性，N-乙酰葡萄糖胺 (NAG)和 N-乙酰甘露糖胺(NAM)交叉排列的共聚物或 N-乙酰葡萄糖胺的聚合物在第 4 和第 5 个糖环之间被切断(图 6-8)。反应时底物的第 4 个 NAM 和酶的D

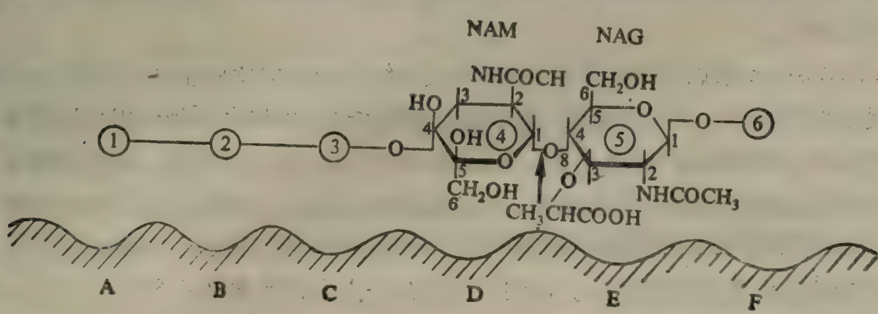


图 6-8 溶菌酶和底物的相对位置

位置结合，糖环从稳定的椅子型变为不稳定的半椅子型，氧原子(O_8)和谷氨酸 35 的 H 成氢键配对，其间距为 0.2 nm。谷氨酸 35 在这里起了一般酶催化剂的作用，限制反应速度的是C_1—O^8间键的切断和正碳离子的生成阶段。

采用推广的 Hückel 法，用图 6-9 所示模型(用甲酸代替谷氨酸 35，二甲氧基甲烷代替底物)，可得如下结果：随 C_1—O_8 键切断的活化能没有酶时为 4.86 eV，有酶时为 4.35 eV。这相当于产生 0.51 eV 稳定能，得到 4×10^8 倍的加速，如图 6-10。

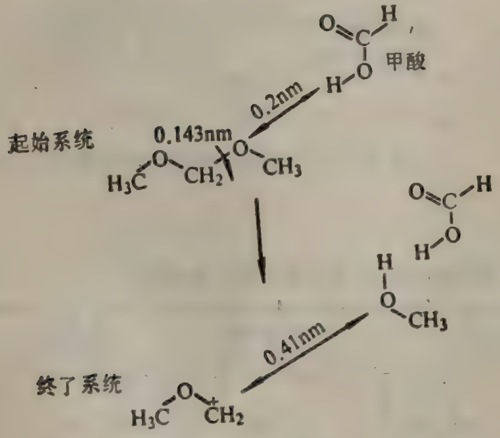


图 6-9 用模型作的溶菌反应机制图
0.143 nm 和 0.41 nm 之间以 0.20，0.225，0.25，0.28，0.31，0.36 nm 的距离计算了能量画出势能

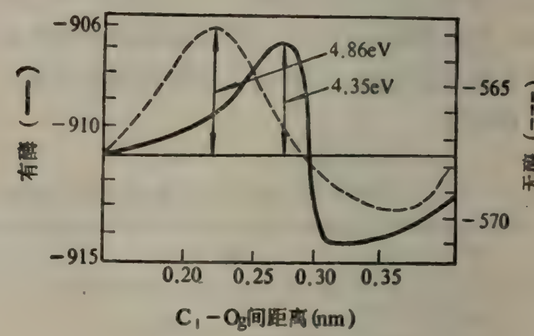


图 6-10 切断糖苷键的位能曲线

Loew 等用推广的 Hückel 方法计算了这个限制速度阶段的活化能，估计了酶的存在对稳定能的作用。计算模型如图 6-9。计算结果表明 C_1—O_8 键切断的活化能没有酶时为 4.86 eV，有酶时为 4.35 eV，产生 0.51 eV 稳定能(图 6-11)，这相当于 4×10^8 倍加速。

图 6-11 是对 α-糜蛋白酶的计算。不考虑有非定域能的计算活化能(虚线表示)较之考虑的情况(实线)要高约2 eV，可以看出在反应的加速中非定域能的作用。

6.1.5 环境的模拟

众所周知，化学反应或多或少总是受环境、其发生处介质的影响，人们已经提出了几种考虑环境影响的理论方案。

早在 1935 年 Stern 强调过在具有极性的酶-底物络合物的空位中能产生偶极，该偶极所造成的静电场是非常重要的。Yomosa 在 1967 年对酶催化过程中静电场的作用作过更明确的讨论。而第一次定量计算则是在 1972 年由 Johannin 和 Kellersohn 完成的。他们所用的点电荷模型是由蛋白质肽基的所有原子所组成的。对许多酶中存在的 α-螺旋也做过此类计算。Hayes 和 Kollman 在对羧肽酶 A 的模型计算时还考虑了侧链原子，他们指出了在酶反应中的靠近催化部位中心机构的带电基团的重要性。他们所用的分子点电荷数据是从 STO-3 G 或 4-31 G 基组中得到的。对硫酸酐酶的计算表明这种分子内静电场对它的反应活性有很大的影响。近来，

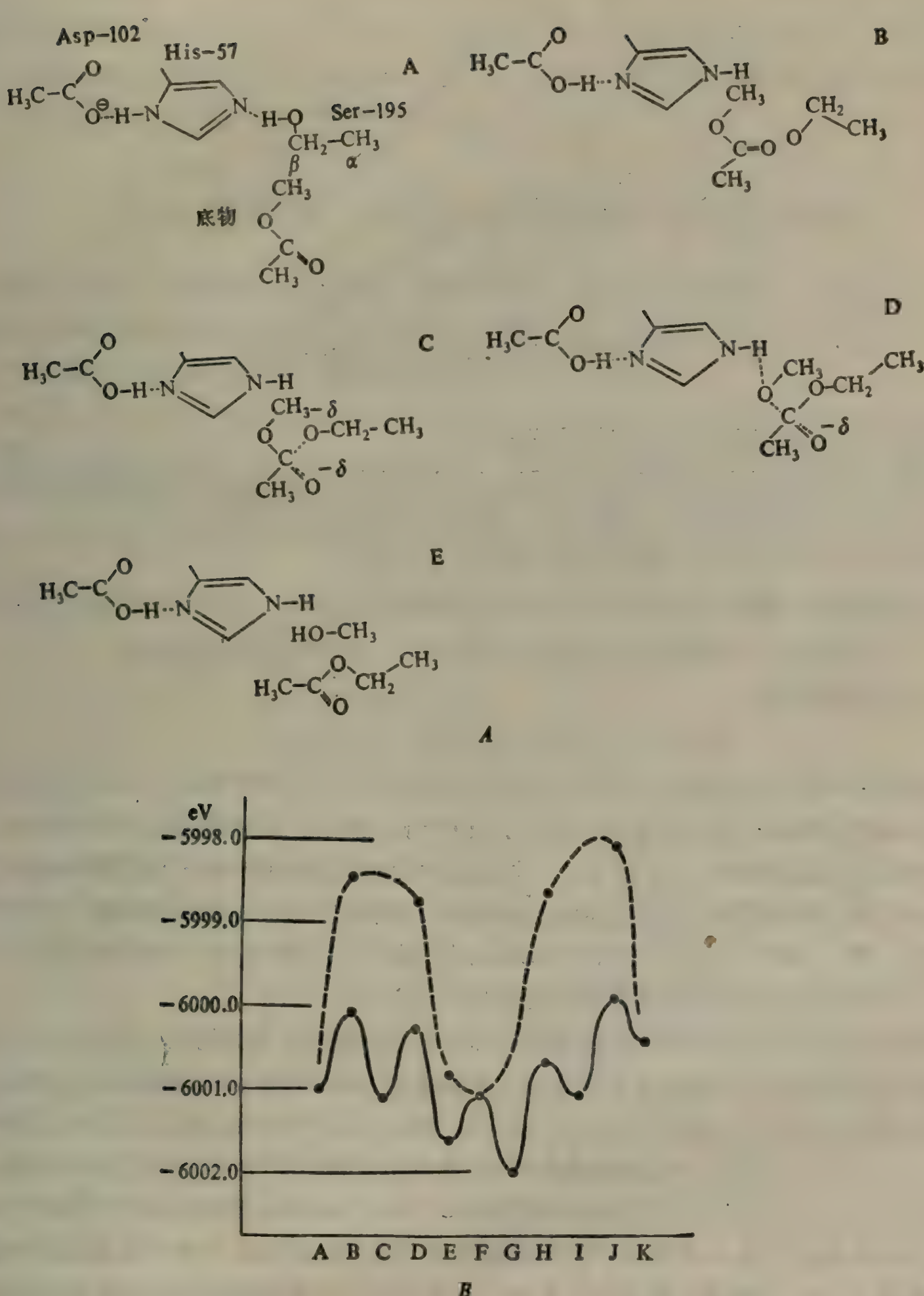


图 6-11 *A*.α-糜蛋白酶反应机制图
B.考虑非定域能(实线)和不考虑(虚线)情况下 α-糜蛋白酶位能曲线

又有人对枯草杆菌蛋白酶的催化三价基的周围静电势作了计算,如图 6-12 所示。另外,在小的模型分子计算基础上,Scrocco 和 Tomasi 将势函数表示为具有可转移性化学键的贡献之和,该

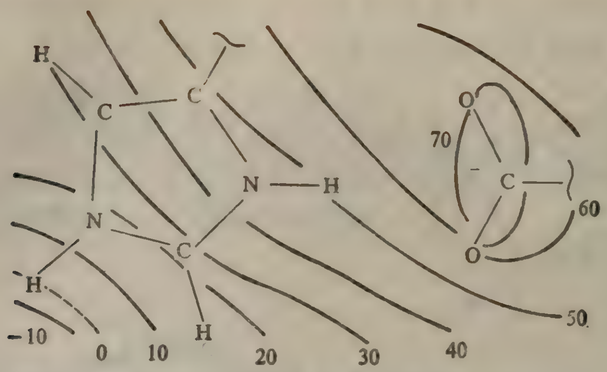


图 6-12 枯草杆菌蛋白酶催化三价基分子静电势(kJ/mol)

方法不仅简化了计算而且能得到很好的计算结果。

为了得到环境对酶反应影响更逼真的描述,可以将整个反应系统分为催化部位和环境两部分。对这两部分用不同的计算方法来处理。反应的总的能量变化可表示为

$$\Delta E = \Delta E_c + \Delta E_{c/env} + \Delta E_{env} \tag{6.1}$$

其中 ΔE_c 代表和催化部位中心机构所对应的气相模型的能量变化,它可用半经验式从头算计算。$\Delta E_{c/env}$ 代表催化中心机构和环境的相互作用;ΔE_{env} 表示环境的能量变化。

$\Delta E_{c/env}$ 又可表示为

$$\Delta E_{c/env} = \varepsilon^{-1} \Sigma [\underline{V}_f(\underline{r}_m) \underline{q}_{fm} - \underline{V}_i(\underline{r}_m) \underline{q}_{im}] \tag{6.2}$$

其中 $V(\underline{r}_m)$表示环境在带电为 q_m 的第 m 个原子所在位置所产生的静电势,下标 $\underline{f}$ 和 $\underline{i}$ 表示终态和始态,ε 为介电常数。若仅考虑第一水化层,由于水和被水化系统间为真空,取值为 1,而对蛋白质骨架来讲,ε 的范围从 1.36 至 4.0。求和包括气相模型的所有原子,它们是从酶中切割出来的,但在对 ΔE_c 进行计算时,断键处由氢原子饱和。例如,在丝氨酸侧链的情形,上述方程中的模型包括—CH_2—OH,但在分子轨道计算时是对甲醇 CH_3—OH 计算。静电势 V 可用点电荷模型得到,也可由其他方法得到,如前面提过的 Scrocco 方法。

更精确的方法是在气相模型的哈密顿量中加上包括环境静电势势能项。该方法相当于考虑由于环境所引起的催化中心机构原子的电荷分布的极化。但该方法没能考虑由催化基团、电荷转移和交换等造成的对蛋白质骨架的反极化作用。Hayes 和 Kollman 指出并非蛋白质核原子的所有部分电荷都是重要的,只有赖氨酸的 NH_3^+ 和天冬氨酸的COO^-以及催化中心机构 1.5nm 以内的那些带电基团是重要的,其余的全可以略去。Van Duijnen 等在讨论木瓜蛋白酶时,只考虑催化中心机构附近的 α-螺旋所产生的静电场的影响。他们指出 α-螺旋造成的影响随螺旋的增长而加强。但当 α-螺旋长度超过 1.0 nm 后,其影响就不再随螺旋长度的增长而加强了。

Tapia 等在气相模型的哈密顿量中不仅包括了环境的静电势项,还加了反应场项。这样就进一步考虑了由气相模型所产生的蛋白质核的极化。他们的研究表明,对肝乙醇脱氢酶的质子转移系统来讲,该项是非常重要的,约有 100～300 kJ/mol。

在研究溶菌酶时,Warshel 和 Levitt 在根据式 6.2 计算 $E_{c/env}$ 时,除了考虑静电效应和极化

项外、还考虑范德华项。计算表明，它们对反应平衡的影响不大，而电荷-电荷，极化相互作用的影响起决定性作用。在计算 ΔE_{tm} 时，他们用的是由 Levitt 提出的经验方法。他们还研究了基态和过渡态构象中的应变效应。Clementi 在研究由木瓜蛋白酶参与的催化反应时，也用了上述方法。现在，ΔE_{env} 的计算已可用分子力学方法进行了。

6.2 羧肽酶 A 和其它金属酶的量子生物学研究

羧肽酶 A 是一个含 Zn^{2+} 的蛋白质水解酶，它能水解蛋白质或多肽的 C 末端肽键，当 C 末端氨基酸残基是芳香性或者具有较大空间的脂肪族侧链时，水解的速度较快，如图 6-13 示。羧肽酶 A 不能水解 C 末端是碱性氨基酸残基和脯氨酸残基的肽键。

羧肽酶 A 的三维空间肽链走向如图 6-14 所示。它的分子是椭园形的，为 5.0×4.2×3.8 nm。其中 α-螺旋占 38%，β-折叠占 17%。Zn^{2+} 和酶紧密地结合在活性部位。活性部位处于酶分子表面的裂缝之中。组氨酸-69 和 169 的咪唑基，谷氨酸-72 的羧基和一个水分子以配位键的方式形成四面体，靠近 Zn^{2+} 部分有一个较大的空间，可以容纳 C 末端氨基酸的侧链，如图 6-15 所示。

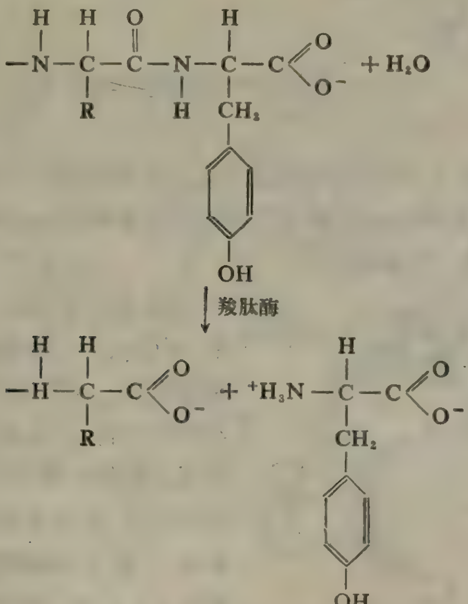


图 6-13 羧肽酶 A 的水解反应

酶和底物是如何结合的？

用甘氨酸和 L-酪氨酸这个底物类似物与酶的晶体，通过 X 射线衍射得到的电子云密度图上，通过理论分析，可以看出其间存在着复杂的相互作用，这种相互作用可以描述如下：(1) 甘氨酸-L-酪氨酸的末端羧基负电荷和精氨酸 145 的侧链胍基形成离子键；(2) 酪氨酸侧链和酶的非极性区通过疏水键结合；(3) 甘氨酸-L-酪氨酸的肽链—NH—作为氢键受体和酪氨酸-248 侧

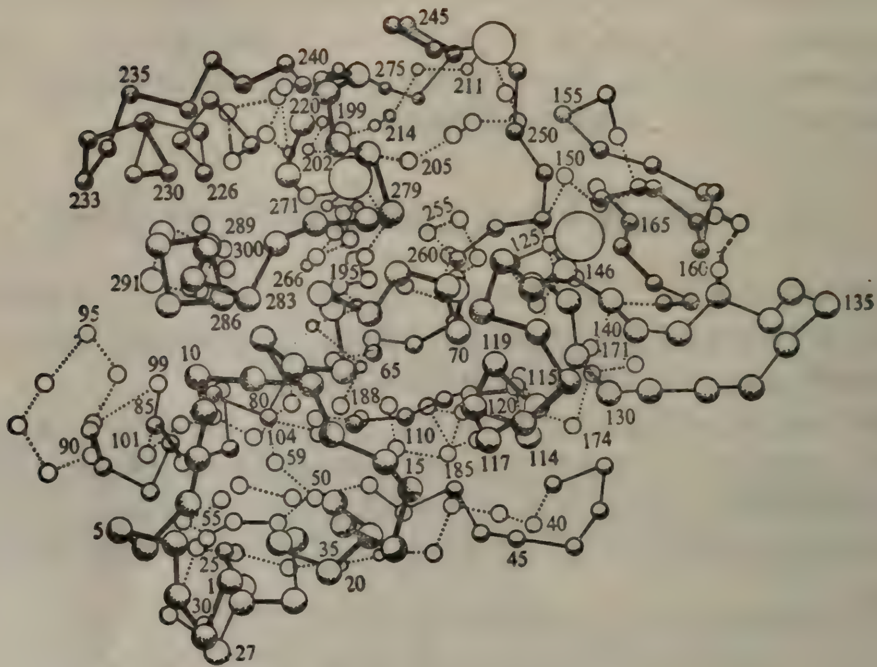


图 6-14 羧肽酶A的三维空间肽链走向

链羧基形成氢键；(4) 甘氨酰-L-酪氨酸的肽链$-\overset{O}{\overset{\parallel}{C}}-$的氧和 Zn^{2+} 形成配位键；(5) N末端氨基和谷氨酸-270 通过一个水分子形成氢键。这种键在正常底物中是不存在的，可能它是使甘氨酸-L-酪氨酸水解速度降低的一个原因。

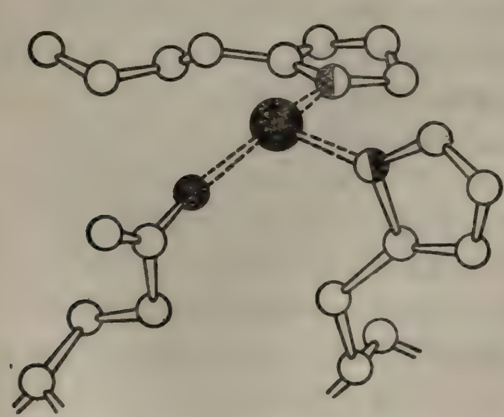


图 6-15 羧肽酶A活性部位 Zn^{2+} 的配位结构

因为甘氨酰-L-酪氨酸能被酶缓缓地水解，可以认为它与酶的结合与正常底物的结合是相似的。比较甘氨酰-L-酪氨酸和酶络合物与酶的结构，可以看出甘氨酰-L-酪氨酸结合以后，引起了整个酶三维结构的较大重排。在活性部位，这种重排使催化基团和被水解的键靠近。可以看到精氨酸-145 和谷氨酸-270 各移动了0.2 nm，原来酶上有一个水分子和 Zn^{2+} 形成配位键，在甘氨酸-L-酪氨酸和酶的络合物中水分子被肽键的氢取代了。非极性口袋中至少有四个水分子被底物类似物的酪氨酸侧链酚基取代。最大的构象变化发生在酪氨酸-248，它的酚基在酶形成络合物时移动了 1.2 nm，大约等于酶分子直径的 1/4，这个移动伴随着 C—C 单键的转动，使酪氨酸-248 的酚基由酶分子的表面移向底物被作用的肽键，见图 6-16。

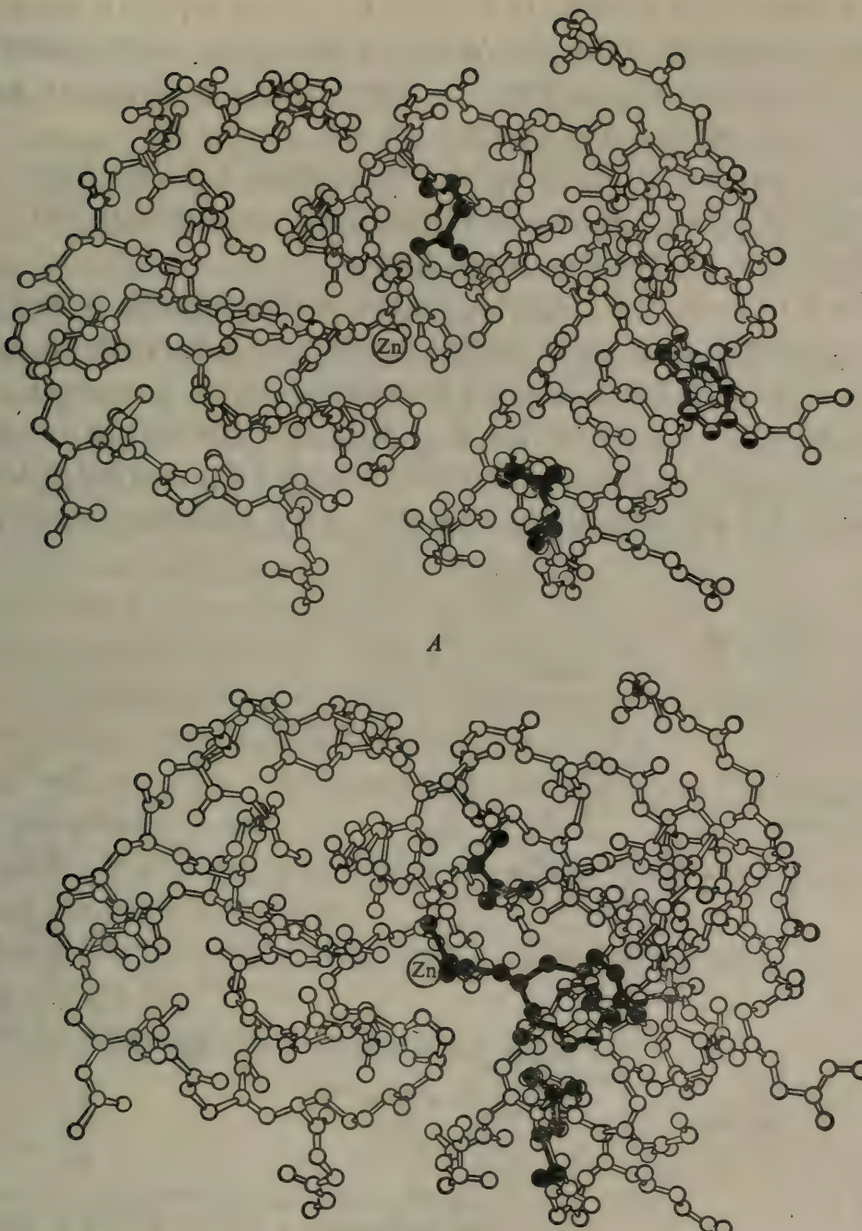


图 6-16 底物结合引起羧肽酶A构象的变化

A. 底物结合前；*B*. 底物结合后。特别注意酪氨酸 248 在底物结合后位置的移动

这种结构的变化是由精氨酸-145 和二肽 C 末端羧基结合开始的，构象变化的结果，使一个原来是充满水的凹洼变成了一个疏水口袋。

酶与底物的相互作用能跟酶的构象有关，对羧肽酶 A，它与甘氨酰-L-酪氨酸的结合能，若

用 STO-3 G 算出的原子净电荷为数据，Zn 为正二价，那么，天然状态的酶与底物的结合能要比结合状态的酶与底物的相互作用能小 10 kJ/mol。这也就说明了为什么酶和底物结合时其构象要发生变化。然而有趣的是，若 Zn 的价数为 1.4（为模拟金属原子正电荷向周围配体转移而造成的非定域现象），则天然状态与底物的结合反而要小些了。

Hayes 和 Kollman 还研究了甘氨酸-酪氨酸类似物与羧肽酶 A 结合的相对能力。他们发现如果苯丙氨酸侧链用邻位取代对位，那么，结合能力则增加，因此可以认为，O—HO—甘氨酸-酪氨酸类似物将比甘氨酸-酪氨酸具有更大的结合能力。

Nakagana 和 Umeyama 研究了羧肽酶 A 与底物结合时精氨酸的作用。在羧肽酶 A 的活性部位中有三个精氨酸残基，精氨酸-71，精氨酸-127 在酶的分子表面，精氨酸-145 在酶的内部，根据 CNDO/2 和从头算，他们提出了底物键合的“滑动模型”(sliding mechanism)。该模型认为底物 C 端的羧基首先被处在最表面的精氨酸-11 所捕获，然而它滑到精氨酸-127，后进入活性部分的精氨酸-145，如图 6-17 所示。他们的计算表明，从能量上来讲，只有羧肽酶 A 与底物结合后，精氨酸-145 和 C 端羧基的相对运动才是有利的。不少作者还用量化方法研究了模型底物

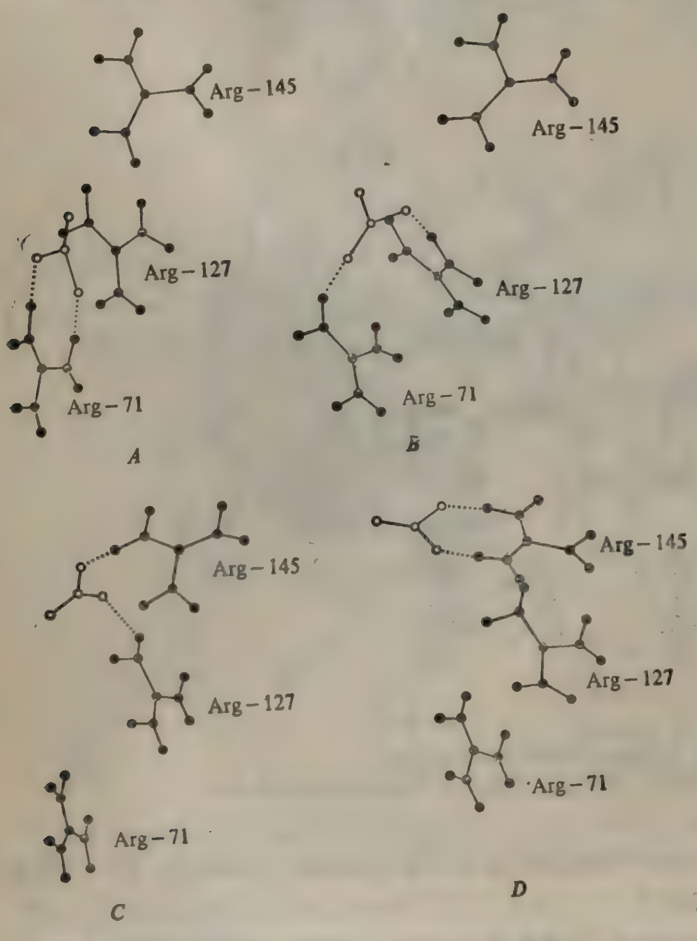


图 6-17 羧肽酶A底物的“滑动模型”。虚线表示氢键

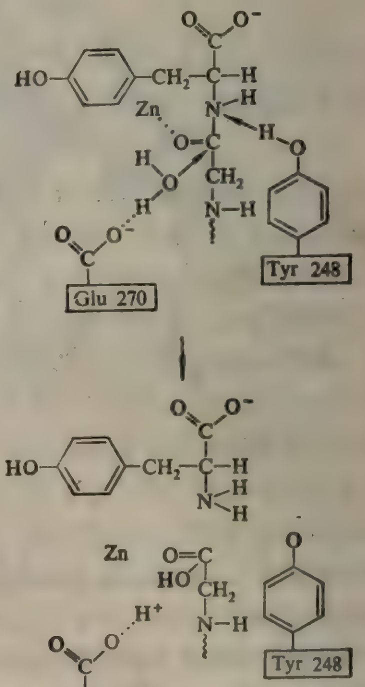


图 6-18 羧肽酶A的作用

中 C—N 集居数的各种变化。

羧肽酶A的正常底物作用机制，根据X射线结晶学的结果和理论研究，可用图 6-18 表示。

谷氨酸-270 的羧基作为广义的碱作用于水分子的氢，通过水分子的 OH 对碳原子进行亲核攻击，Zn^{2+} 起着酸的功能，和羧基的氧作用，使得电子由碳原子向氧原子方向流动，造成羧基碳原子上电子缺少，易于受到 OH 的亲核攻击。酪氨酸-248 和氮原子形成的氢键，进一步使氮原子 sp^3 杂化，于是碳原子离开了原来的酰胺平面，有利于肽链的水解。在酶和甘氨酸—L-酪氨酸二肽结合时，由于谷氨酸-270 羧基和二肽的末端氨基形成氢键，失去了它活化水分子对二肽肽键羧基碳原子亲核进攻的能力，所以二肽的水解速度很慢。

金属离子在催化中的作用，仅仅用实验是很难确定的，因此理论研究就显得特别重要了。人们发现金属离子 Zn^{2+} 的正电荷在增加羧基 C 原子上的净电荷和稳定由组氨酸-69 和- 196 的咪唑基，谷氨酸-72 羧基和一个水分子形成的四面体中有其重要的作用。Sheiner 和 Lipscomb 用 Li^{-1} 和 Be^{2+} 来取代 Zn^{2+} 得到了同样的结论，他们指出，四面体的稳定是由于金属亲电性和羧基氧间的部分成键所造成。并且金属亲电性对肽键水解的影响比氢键亲电性的 NH_4^+ 和OH_2要强。Dsman 和 Weinstern 还指出，Zn^{2+} 的作用比 Be^{2+} 强烈。

综上所述，羧肽酶A的作用有如下几点：

(1) 羧肽酶A和底物结合后，有一个较大的三维结构变化；

(2) 这个构象的变化很可能是由底物的C末端羧基和精氨酸-145 侧链胍基的结合。通过“滑动机制”而开始的，因此底物如果没有羧基，也就不能水解。

(3) 这个构象的变化造成了催化基因与被催化水解肽键的相互接近，同时形成了一个疏水口袋、底物与酶进一步结合；

(4) 肽键的水解通过亲核反应进行，由谷氨酸-270 作为广义碱通过水分子进行亲核进攻’这一作用为酪氨酸-248 和 Zn^{2+} 所加强。

(5) Zn^{2+} 在整个催化过程中起了增加 C 原子上正电荷（有利用亲核进攻）和稳定四面体（有利疏水口袋）的重要作用。

此外，其它金属酶的量子生物学研究请参阅第十章的有关内容。

参考文献

[1] 江寿平译，《量子生物学入门》，上海科学技术出版社。
[2] 刘培南等，《基础分子生物学》，高等教育出版社，1983。
[3] 许根俊，《酶的作用原理》，科学出版社，1983。
[4] 吴加金，《生物化学与生物物理进展》，第1期，1983年。
[5] I. G. Csizmadia, Molecular Structure and Conformation, Amsterdam-Oxford-New York(1982).
[6] Szaao, G. Naray, et al Inter. J. Quant. Chem., 7, 397,(1980).
[7] W. N. Lipscomb, Anu. Rev. Biochem., 52,17(1983).
[8] R. B. Brant, J. Laux, C. Thomson et al, Int. J. Quantum. Chem., QBll (1984).
[9] O. Tapia, C. Lamborelle and G. Johannin, Chem. Phys. Letters 72 (1980) 334.
[10] A. Warshel and M. Levitt, J. Mol. Biol. 103 (1976) 227.
[11] H. Umeyama, A. Imamura, C. Nagata and M. Hanano, J. Theor. Biol. 41 (1973) 485.

[12] A. Warshel, Proc. Natl. Acad. Sci. USA 75 (1978) 5250.
[13] D. M. Hayes and P. Kollman, in B. Pullman (Ed.), 12th Jerusalem Symposium: Catalysis in Chemistry and Biochemistry, Reidel, Dordrecht, 1980, p. 77.
[14] G. Náray-Szabó, in F. Darvas (Ed.), Proc. 3th Congr. Hungarian Pharm. Soc. : Quantitative Structure-Activity Relationships, Akadémiai Kiadó, Budapest, 1980, p. 295.
[15] G. Náray-Szabó and L. Polgár, Int. J. Quant. Chem. Quantum Biol. Symp. 7 (1980) 397.
[16] O. Tapia in, R. Daudel, A. Pullman, L. Salem and A. Veillard (Eds.), Quantum Theory of Chemical Reactivity, Vol. 2, Reidel, Dordrecht, 1980, p. 220
[17] P. T. Van Duijnen, B. T. Thole and W. G. J. Hol, Biophys. Chem. 9 (1979) 273.
[18] R. Ditchfield, W. J. Hehre and J. A. Pople, J. Chem. Phys. 54 (1971) 724.
[19] S. Nakagawa, H. Umeyama and T. Kudo, Chem. Pharm. Bull. 28 (1980) 1342.
[20] R. Osman and H. Weinstein, Isr. J. Chem. 19 (1980) 149.
[21] W. J. Hehre, R. F. Stewart and J. A. Pople, J. Chem. Phys. 51 (1969) 2657.
[22] R. Lavery and B. Pullman, Int. J. Quant. Chem. Quantum Biol. Symp. 6 (1979) 467.
[23] T. A. Halgren and W. N. Lipscomb, J. Chem. Phys. 58 (1973) 1569
[24] S. Scheiner, D. A. Kleier and W. N. Lipscomb, Proc. Natl. Acad. Sci. USA 72 (1975) 2606.
[25] J. C. Slater in, P. -O. Löwdin (Ed.), Advances in Quantum Chemistry, Vol. 6, Academic Press, New York, 1972, p. 1.
[26] P. J. M. Geurts, J. W. Gosselink, A. van der Avoird, E. J. Baerends and J. G. Snijders, Chem. Phys. 46 (1980) 133.
[27] R. C. Bingham, M. J. S. Dewar and D. H. Lo, J. Am. Chem. Soc. 97 (1975) 1285.
[28] A. T. Pudzianowski and G. H. Loew, J. Am. Chem. Soc. 102 (1980) 5443.
[29] J. A. Pople and D. L. Beveridge, Approximate Molecular Orbital Theory, McGraw Hill, New York, 1970.
[30] H. P. Kitayama and H. Fukutome, J. Theor. Biol. 60 (1976)1.
[31] R. Hoffman, J. Chem. Phys. 39 (1963) 1397.
[32] L. K. Hanson, Int. J. Quant. Chem. Quantum Biol. Symp. 6 (1979) 73.
[33] P. Pulay in, H. F. Schaefer III (Ed.), Modern Theoretical Chemistry, Vol. IV : Applications of Electronic Structure Theory, Plenum, New York, 1977, p. 153.
[34] P. Čársky and M. Urban, Ab Initio Calculations· Methods and Applications in Chemistry, Lecture Notes in Chemistry, Vol. 16, Springer Berlin, 1980.
[35] P. Schuster in, P. Schuster, G. Zundel and C Sandorfy (Eds.), The Hydrogen Bond, Vol. I : Theory, North-Holland, Amsterdam, 1976.
[36] H. F. Schaefer III (Ed.), Modern Theoretical Chemistry, Vol. IV: Applications of Electronic Structure Theory, Plenum, New York, 1977
[37] G. A. Segal (Ed.), Modern Theoretical Chemistry, Vol. VI: Semiempilical Molecular Orbital Theory: Applications, Plenum, New York, 1977.
[38] R. Yamdagni and J. Kebarle, J. Am. Chem. Soc. 95 (1975) 4050.
[39] G. Náray-Szabó, A. Kapur, P. G. Mezey and L. Polgár, J. Am.Chem. Soc. Submitted for publication.
[40] (a) M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. 99 (1977) 4899;
(b) M. J. S. Dewar and M. L. Mckee, J. Am. Chem. Soc. 99 (1977) 5231; (C) M. J. S. Dewar and H. J. Rzepa, J. Am. Chem. Soc. 100 (1978) 58.
[41] W. Thiel, J. Am. Chem. Soc. 103 (1981) 1413.
[42] G. R. Pack and G. H. Loew, Int. J. Quant. Chem. Quantum Biol. Symp. 6 (1979) 381.
[43] P. R. Andrews, G. D. Smith and I. G. Young, Biochemistry 12 (1973) 3492.
[44] L. L. Ingraham, Biochim. Biophys. Acta 304 (1973) 891.
[45] Y. Beppu and S. Yomosa, J. Phys. Soc. Jpn. 42 (1977) 1694.
[46] Y. Beppu and S. Yomosa, J. Phys. Soc. Jpn. 44 (1978) 291.
[47] S. Nakagawa and H. Umeyama, Chem. Pharm. Bull. 25 (1977) 909.

[48] G. I. Amidon, J. Theor. Biol. 46 (1974) 101.
[49] H. Umeyama, A. Imamura and C. Nagata, Chem. Pharm. Bull. 23 (1975) 3045.
[50] H. Umeyama, Chem. Pharm. Bull. 22 (1974) 2518.
[51] G. H. Loew, L. M. Hjelmeland and R. F. Kirchner, Int. J. Quant. Chem. Quantum Biol. Symp. 4 (1977) 225.
[52] H. Umeyama, Chem. Pharm. Bull. 28 (1980) 1317.
[53] M. M. Rohmer and G. H. Loew, Int. J. Quantm. Chem. Quantum Biol. Symp. 6 (1979) 93.
[54] S. Scheiner and W. N. Lipscomb, Proc. Natl. Acad. Sci. USA 73 (1976) 432.
[55] H. Umeyama and S. Nakagawa, Chem. Pharm. Bull. 27 (1979) 1524.
[56] R. Broer, P. Th. Van Duijnen and W. C. Nieuwpoort, Chem. Phys. Letters 42 (1976) 525.
[57] J. A. Pople in, H. F. Schaefer III (Ed.), Modern Theoretical Chemistry, Vol. IV: Applications of Electronic Structure Theory, Plenum, New York, 1977, p. 1.
[58] F. Török, A. Hegedüs and P. Pulay, Theoret. chim. Acta (Berl.) 32 (1973) 145.
[59] N. L. Allinger, Advan. Phys. Org. Chem. 13 (1976)1.
[60] G. Némethy and H. A. Scheraga, Quart. Rev. Biophys. 10 (1977) 239.
[61] A. Warshel in, G. A. Segal (Ed.), Modern Theoretical Chemistry, Vol. VI : Semiempirical Molecular Orbital Theory: Applications, Plenum, New York, 1977, p. 133.
[62] F. A. Momany, R. F. McGuire, A. W. Burgess and H. A. Scheraga, J. Phys. Chem. 79 (1975) 2361.
[63] E. Clementi, F. Cavallone and R. Scordamaglia, J. Am. Chem. Soc. 99 (1977) 5531.
[64] I. Tvaroska and T. Bleha, Biopolymers 18 (1979) 2537.
[65] M. R. Pincus, A. W. Burgess and H. A. Scheraga, Biopolymers 15 (1976) 2485.
[66] A. T. Hagler, E. Huler and S. Lifson, J. Am. Chem. Soc. 96 (1974) 5319.
[67] A. T. Hagler, P. S. Stern, R. Sharon, J. M. Becker and F. Naider, J. Am. Chem. Soc. 101 (1979) 6842.
[68] F. A. Momany, L. M. Carruthers, R. F. McGuire and H. A. Scheraga, J. Phys. Chem. 78 (1974) 1595.
[69] M. R. Pincus, S. S. Zimmerman and H. A. Scheraga, Proc. Natl. Acad. Sci. USA, 74(1977)4261.
[70] D. F. DeTar, J. Am. Chem. Soc. 103 (1981) 107.
[71] M. R. Pincus, S. S. Zimmerman and H. A. Scheraga, Proc. Natl. Acad. Sci. USA 74 (1977) 2629.
[72] M. Levitt in, E. R. Blout, F. A. Bovey, M. Goodman and N. Lotan (Eds.), Peptides, Polypeptides and Proteins, Wiley, New York, 1974, p. 99.
[73] E. Clementi, G. Ranghino and R. Scordamaglia, Chem. Phys. Letters 49 (1977) 218.
[74] A. T. Hagler and J. Moult, Nature 272 (1978) 222.
[75] J. Hermans, D. R. Ferro, J. E. McQueen and S. C. Wei in, B. Pullman (Ed.), Environmental Effects on Molecular Structure and Properties, Reidel, Dordrecht, 1976. p. 459.
[76] P. Cremaschi, A. Gamba and M. Simonetta, Theoret. chim. Acta (Berl.) 25 (1972) 237.
[77] S. Yamabe, S. Kato, H. Fujimoto and K. Fukui, Theoret. chim. Acta (Berl.) 30 (1973) 327.
[78] O. Tapia, E. Poulain and F. Sussman, Chem. Phys. Letters 33 (1975) 65.
[79] G. Náray-Szabó, P. Pulay and P. Mezey, Acta chim. Acad. Sci. Hung. 90 (1976) 199.
[80] O. Tapia and E. Poulain, Int. J. Quant. Chem. 11 (1977) 473.
[81] O. Tapia and J. E. Sanhueza, Biochem. Biophys. Res. Commun. 81 (1978) 336.
[82] M. Taagepera, D. Defrees, W. J. Hehre and R. W. Taft, J. Am. Chem. Soc. 102 (1980) 424.
[83] A. Gamba, M. Simonetta, G. Safrittl, I. Szele and H. Zollinger, JCS Perkin II (1980) 493.
[84] Stern, quoted by R. Lumry in, P. D. Boyer. H. Lardy and K. Myrbäck (Eds.), The Enzymes 2nd Ed., Vol. I, Academic, New York, 1959, p. 157, ref. 166.
[85] S. Yomosa, Progr. Theoret. Phys. Suppl. 40 (1967) 249.
[86] G. Johannin and N. Kellersohn, Biochem. Biophys. Res. Commun. 49 (1972) 321.
[87] R. P. Sheridan and L. C. Allen, Biophys. Chem. 11 (1980) 133.

[88] W. G. J. Hol, P. Th. Van Duijnen and H. J. C. Berendsen, Nature 273 (1978) 443.
[89] D. M. Hayes and P. A. Kollman, J. Am. Chem. Soc. 98 (1976) 3335.
[90] D. M. Hayes and P. A. Kollman, J. Am. Chem. Soc. 98 (1976) 7811.
[91] R. P. Sheridan and L. C. Allen, J. Am. Chem. Soc. 103 (1981) 1544.
[92] E. Scrocco and J. Tomasi, Top. Current Chem. 42 (1973) 95.
[93] G. Náray-Szabó, Int. J. Quant. Chem. 16 (1979) 265.
[94] G. Náray-Szabó, A. Grofcsik, K. Kósa, M. Kubinyi and A. Martin, J. Comput. Chem. 2 (1981) 58.
[95] G. Náray-Szabó, Quantum Chemistry Program Exchange 13 (1980) 396.
[96] G. Náray-Szabó, unpublished.
[97] O. Tapia, F. Sussman and E. Poulain, J. Theor. Biol. 71 (1978) 49.
[98] P. Th. Van Duijnen, B. T. Thole, R. Broer and W. C. Nieuwpoort, Int. J. Quant. Chem. 17 (1980) 651.
[99] M. Levitt, J. Mol. Biol. 82 (1974) 393.
[100] P. Claverie, J. -P. Daudey,J. Langlet, B. Pullman, D. Piazzola and M. J. Huron, J. Phys. Chem. 82 (1978) 405.
[101] B. Pullman, S. Miertuš and D. Perahia, Theoret. Chim. Acta (Berl.) 50 (1978) 317.
[102] K. Jug, Theoret. chim. Acta (Berl.) 54 (1980) 263.
[103] J. P. Klinman in, R. D. Gandour and R. L. Schowen (Eds.), Transition States of Biochemical Processes, Plenum, New York-London, 1978, p. 165.
[104] M. D. Harmony, Chem. Soc. Rev. 1 (1972) 211.
[105] P. Baňacký, Biophys. Chem. 13 (1981) 39.
[106] P. Hobza, P. čársky and R. Zahradník, Int. J. Quant. Chem. 16 (1979) 257
[107] H. A. Scheraga, Adv. Phys. Org. Chem. 6 (1968) 103.
[108] J. W. Larsen, Biochem. Biophys. Res. Commun. 50 (1973) 839.
[109] C. Delisi and D. M. Crothers, Biopolymers 12 (1973) 1689.
[110] J. Achimowitz, A. Cader, L. Pannert and E. Wojcik, Phys. Letters 60A (1977) 383.
[111] M. V. Volkenstein, I. B. Golovanov and V. M. Sobolev, Int. J. Quant. Chem. 16 (1979) 777.
[112] B. Somogyi and S. Damjanovich, J. Theor. Biol. 51 (1975) 393.
[113] A. Cooper, Proc. Natl. Acad. Sci. USA 73 (1976) 2740.
[114] L. B. Kier, Molecular Orbital Theory in Drug Research. Academic, New York, 1971.
[115] L. B. Kier and H. -D. Höltje, J. Theor. Biol. 49 (1975) 401.
[116] D. M. Blow, J. J. Birktoft and B. S. Hartley, Nature 221 (1969) 337.
[117] L. Polgár and M. L. Bender, Proc. Natl. Acad. Sci. USA 64 (1969) 1335.
[118] W. W. Bachovchin and J. D. Roberts, J. Am. Chem. Soc. 100 (1978) 8041.
[119] J. L. Markley and I. B. Ibanez, Biochemistry 17 (1978) 4627.
[120] A. A. Kossiakoff and S. A. Spencer, Nature 288 (1980) 414.
[121] G. Náray-Szabó and J. Angyán, in preparation.
[122] S. E. Hamilton and B. Zerner, J. Am. Chem. Soc. 103 (1981) 1827.
[123] M. W. Hunkapiller, S. H. Smallcombe, D. R. Whitaker and J. H. Richards, Biochemistry 12 (1973) 4732.
[124] P. A. Kollman and I. D. Kuntz, J. Am. Chem. Soc. 98 (1976) 5820.
[125] J. J. Birktoft and D. M. Blow, J. Mol. Biol. 68 (1972) 187.
[126] D. A. Matthews, R. A. Alden, J. J. Birktoft, S. T. Freer and J. Kraut, J. Biol. Chem. 252 (1977) 8875.
[127] R. A. Alden, J. J. Birktoft, J. Kraut, J. D. Robertus and C. S. Wright, Biochem. Biophys. Res. Commun. 45 (1971) 337.
[128] J. Drenth, J. N. Jansonius, R. Koekoek and B. G. Wolthers, Adv. Protein Chem. 25 (1971) 79.

[129] J. Drenth, K. H. Kalk and H. M. Swen, Biochemistry 15 (1976) 3731.
[130] R. Broer, P. Th. Van Duijnen and W. C. Nieuwpoort, Chem. Phys. Letters 42 (1976) 525.
[131] (a) K. Brocklehurst and G. Little, Biochem. J. 128 (1972) 471; (b) A. Williams, E. C. Lucas and K. T. Douglas, JCS Perkin II (1972) 1493; (c) p. Campbell and E. T. Kaiser, J. Am. Chem. Soc. 95 (1973) 3735.
[132] (a) L. Polgár, Eur. J. Biochem. 33 (1973) 104; (b) L. Polgár, FEBS Letters 38 (1974) 187; (c) G. Lowe and A. S. Whitworth, Biochem. J. 141 (1974) 503.
[133] P. Th. Van Duijnen, B. T. Thole, R. Broer and W. C. Nieuwpoort, Int. J. Quant. Chem. 17(1980) 651.
[134] C. C. F. Blake, L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips and V. P. Sharma, Proc. Roy. Soc. (London) B167 (1967) 378.
[135] B. M. Dunn and T. C. Bruice, Adv. Enzymol. 37 (1973) 1.
[136] M. R. Pincus and H. A. Scheraga, Macromolecules 12 (1979) 633.
[137] G. H. Loew and D. D. Thomas, J. Theor. Biol. 36 (1972) 89
[138] M. Zerner, M. Gouterman and H. Kobayashi, Theoret. chim. Acta (Berl.) 6 (1966) 363.
[139] D. C. Gorenstein, J. B. Findlay, B. A. Luxon and D. Kar, J. Am. Chem. Soc. 99 (1977) 3473.
[140] I. Tvaroška and T. Bleha, Can. J. Chem. 57 (1979) 424.
[141] D. M. Chipman and N. Sharon, Science 165 (1969) 454.
[142] L. O. Ford, L. N. Johnson, P. A. Machin, D. C. Phillips and R. Tjian, J. Mol. Biol. 88 (1974) 349.
[143] I. B. Golovanov, V. M. Sobolev, E. V. Shukina and M. V. Volkenstein, Mol. Biol. (Moscow) 10 (1976) 1103.
[144] V. Klimo, I. Tvaroška and T. Bleha, Biophys. Struct. Mech. submitted for publication.
[145] (a) A. Warshel and R. Weiss, J. Am. Chem. Soc. 102 (1980) 6218;
(b) A. Warshel, Biochemistry 20 (1981) ××××.
[146] F. A. Quiocho and W. N. Lipscomb, Adv. Protein Chem. 25 (1971) 1.
[147] (a) R. Breslow and D. E. McClure, J. Am. Chem. Soc. 98 (1976) 258; (b) R. Breslow and D. Wernick, J. Am. Chem. Soc. 98 (1976) 259.
[148] W. N. Lipscomb, J. A. Hartsuck, F. Quiocho, P. H. Bethge, M. L. Ludwig, T. A. Steitz, H. Muirhead and J. C. Coppola, Brookhaven Symp. Biol. 21 (1968) 24.
[149] S. Nakagawa and H. Umeyama, J. Am. Chem. Soc. 100 (1978) 7716.
[150] S. Scheiner and W. N. Lipscomb, J. Am. Chem. Soc. 99 (1977) 3466.
[151] Y. Pocker and S. Sarkanen, Adv. Enzymol. Relat. Areas Mol. Biol. 47 (1978) 149.
[152] K. K. Kannan, A. Liljas, I. Vaara, P. C. Bergsten, S. Lövgren, B. Strandberg, U. Bengtsson, U. Carlborn, K. Fridborg, L. Jarup and M. Petef, Cold Spring Harbor Symp. Quantum Biol. 36 (1971) 221.
[153] K. K. Kannan, B. Notstrand, K. Fridborg, S. Löwgren, A. Ohlsson and M. Petef, Proc. Natl. Acad. Sci. USA 72 (1975) 51.
[154] B. Notstrand, I. Vaara and K. K. Kannan in, C. L. Markert (Ed.), Isozymes I, Academic New York, 1975. p. 575.
[155] J. M. B. Harrowfield, V. Morris and A. M. Sargeson, J. Am. Chem. Soc. 98 (1976) 7282.
[156] J. M. Pesando, Biochemistry 14 (1975) 681.
[157] D. Demoulin and A. Pullman, Theoret. chim. Acta (Berl.) 49 (1978) 161.
[158] A. Pullman and D. Demoulin, Int. J. Quant. Chem. 16 (1979) 641.
[159] A. Sawaryn and A. Sokalski, Int. J. Quant. Chem. 16 (1979) 293.
[160] Y. Jean and J. -M. Lehn, Chem. Phys. 39 (1979) 111.
[161] E. Clementi, G. Corongiu, B. Jönsson and S. Romano, Gazz. Chim. Ital. 109 (1979) 669.
[162] D. Demoulin, A. Pullman and B. Sarkar, J. Am. Chem. Soc. 99 (1977) 8498.
[163] K. K. Kannan, M. Petef, K. Fridborg, H. Cid-Dresdner and S. Lövgren, FEBS Letters 73

(1977) 115.
[164] E. Clementi, G. Corongiu, B. Jönsson and S. Romano, FEBS Letters, 100 (1979) 313.
[165] C. -I. Brändén, H. Jörnwall, H. Eklund and B. Furugren in, P. D. Boyer (Ed.). The Enzymes. 3rd Ed., Vol. 3, Academic, New York, 1975, p. 323.
[166] R. R. Holmes, J. A. Deiters and J. C. Gallucci, J. Am. Chem. Soc. 100 (1978) 7393.
[167] R. R. Holmes, Int. J. Pept. Protein Res. 8 (1976) 445.
[168] G. C. K. Roberts, E. A. Dennis, D. H. Meadows, J. S. Cohen and O. Jardetzky, Proc. Natl. Acad. Sci. USA 62 (1969) 1151.
[169] C. A. Deakyne and L. C. Allen, J. Am. Chem. Soc. 101 (1979) 3951.
[170] R. F. Kirchner and G. H. Loew, J. Am. Chem. Soc. 99 (1977) 4339.
[171] I. B. Golovanov, V. M. Sobolev and M. V. Volkenstein in, Fizika Molekul, Vol. 3, Naukova Dumka, Kiev, 1976, p. 80.
[172] L. Turi-Nagy, P. Pelikán, M. Liška, M. Haring, M. Čeppan and M. Breza, Int. J. Quant. Chem. 16 (1979) 485.
[173] P. P. Andrews and P. C. Haddon, Aust. J. Chem. 32 (1979) 1921.
[174] Ya. Kaneti, J. Theor. Biol. 38 (1973) 169.
[175] C. Nagata and T. Yamagouchi, J. Med. Chem. 22 (1979) 13.
[176] H. Umeyama, A. Imamura and C. Nagata, J. Theor. Biol. 46 (1974) 1.
[177] L. Salem, O. Eisenstein, N. T. Anh, H. B. Bürgi, A. Devaquet, G. Segal and A. Veillard, Nuov. J. Chim. 1 (1977) 335.
[178] A. R. Fersht in, P. Dunnill, A. Wiseman and N. Blakebrough (Eds.), Ellis Horwood, Chichester, 1980, p. 13.
[179] M. I. Page and W. P. Jencks, Proc. Natl. Acad. Sci. USA 68 (1971) 1678.
[180] See *e. g.* K. S. Pitzer and L. Brewer, Thermodynamics, McGraw Hill, New York, 1961.
[181] D. R. Storm and D. E. Koshland Jr., Proc. Natl. Acad. Sci.USA 66 (1970) 445.
[182] G. A. Dafforn and D. E. Koshland Jr., Bioorg. Chem. 1 (1971) 129.
[183] G. A. Dafforn and D. E. Koshland Jr., Proc. Natl. Acad. Sci. USA 68 (1971) 2463.
[184] W. P. Jencks and M. I. Page. Biochem. Biophys. Res. Commun. 57 (1974) 887.
[185] S. Scheiner, W. N. Lipscomb and D. A. kleier, J. Am. Chem. Soc. 98 (1976) 4770.
[186] D. G. Hoare, Nature 236 (1972) 437.
[187] R. D. Gandour in, R. D. Gandour and R. L. Schowen (Eds.), Transition States of Biochemical Processes, New York-London, 1978. p. 529.
[188] M. F. Aldersley, A. J. Kirby, P. W. Lancaster, R. S. McDonald and C. P. Smith, JCS Perkin II (1974) 1487.
[189] A. Warshel, Proc. Natl. Acad. Sci. USA 74 (1977) 1789.
[190] A. Warshel and R. M. Weiss, J. Am. Chem. Soc. 103 (1981) 446.
[191] M. Levitt, Ph. D. Thesis, University of Cambridge, 1972.
[192] C. Nagata, H. Fujita and A. Imamura, Chem. -Biol. Interact. 12 (1976) 1.

第 7 章　量子遗传学

7.1　量子遗传学的背景

7.1.1　从经典遗传学到量子遗传学[1-5]

遗传和变异是最基本的生命现象之一。遗传学主要是研究生物怎样把性状传递给后代和生物在传代中怎样发生变异，即生物的信息怎样从细胞传递至细胞，从亲代传至子代，从而得以代代相传；除此而外，还有生物信息在传递过程中的失真及其机制等。

在 Darwin 发表了著名的"物种起源"之后不久，经典遗传学的奠基人奥地利学者 Mendel根据连续八年豌豆杂交实验结果，在 1865 年提出了两条著名的遗传规律：分离规律和自由组合规律。Mendel 的杂交实验是用豌豆的七对形态上的相对性状来进行的。例如红花对白花、高茎对矮茎、种子光滑对皱皮等等。Mendel 发现在子一代中，每对相对性状都只有显性性状得到表现，如子一代全是红花没有白花；而在子二代中会出现隐性性状；显、隐二种性状之比为 3:1，如 3/4 红花，1/4 白花。据此 Mendel提出"遗传因子"的概念来解释上述实验结果，如图 7-1。由此总结出一条分离规律：遗传因子在体细胞中成对存在，互不影响，而当机体形成配子时，成对的遗传因子彼此分离而分配于配子；每个配子只有其中一个因子。Mendel 又比较了两对性状的杂交实验：黄色光滑种子×绿色皱皮种子，发现子一代种子全部为黄色光滑，而子二代种子有黄色光滑、黄色皱皮、绿色光滑、绿色皱皮四种类型，其数量比分别为 9/16、3/16、3/16、1/16，如图 7-2。由此他提出了自由组合规律：两对以上的遗传因子分离时是独立的，在配子中的分配是随机的。但是 Mendel 学说的遭遇不象 Darwin 学说，在当时没有引起人们的重视。直到 1900 年，Mendel 的文章才被三个植物学家荷兰的 H. de Vries、德国的 Correns 及奥地利的 Von Tschermak Seysenegg在同一年里分别发现。他们各自独立地在自己的研究中获得了 Mendel 规律的证据。此后 Mendel 的遗传因子假说才开始引起生物学界的普遍重视。1905 年，英人 Bateson 将此新发展的学科命名为"遗传学"(genetics)；1909 年，Johannsen 将 Mendel 的遗传因子称为"基因"(gene)。

细胞遗传学的基础是 Mendel 遗传学和细胞学。它主要是从染色体的结构和行为来研究遗传现象、机理及其规律。早在 1883 年，Roux 就已推测细胞核内之染色体携带着遗传因子。1903 年 Sutton 研究了昆虫精子的产生，发现染色体两两成对，在减数分裂中的行为同 Mendel 遗传因子对的传递方式完全平行，于是提出 Mendel 遗传因子在染色体上，并设想在一条染色体上存在着多个因子。同年，Boveri 也发现了这一平行关系，Sutton-Boveri 假说成为遗传的染色体理论之萌芽，亦即细胞遗传学的先声。1901 年 Morgan 发表"果蝇性连锁遗传"一文，Morgan 及其学派通过以果蝇眼色这一性状作杂交实验，发现了伴性遗传。连锁实验是用两对相对性状：黑、灰身和长、残翅进行，即灰身(V 显性)残翅(g 隐性)×黑身(v 隐性)长翅(G 显性)，结果是子一

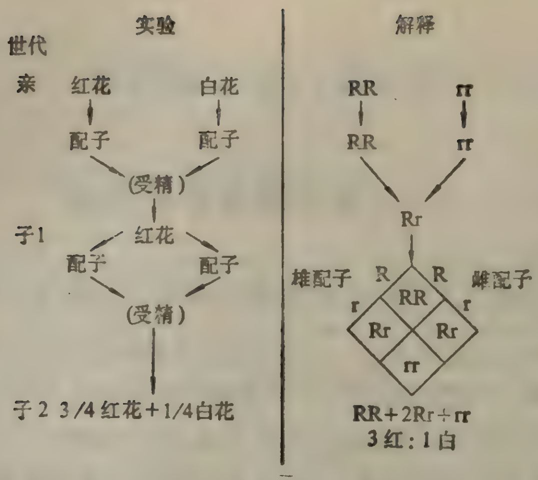


图 7-1 分离规律的实验和解释

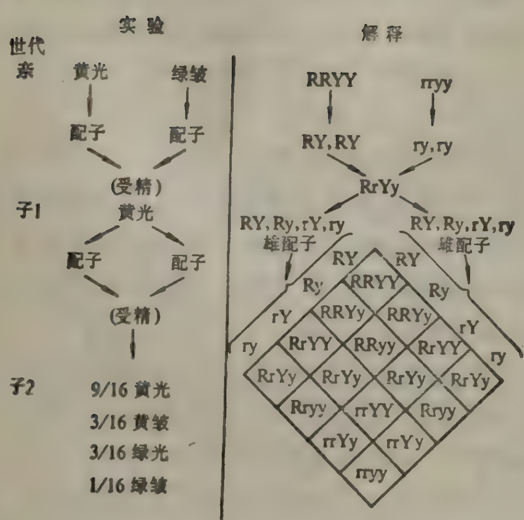
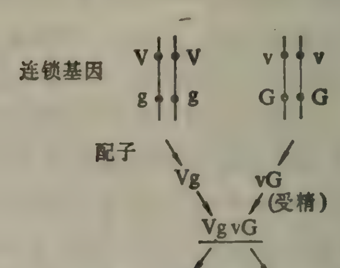


图 7-2 自由组合规律的实验和解释



子1代配子	Vg	vG	
子2代	Vgvg	vGvg	vg 双隐性亲本配子
比例	50%	50%	

图 7-3 连锁假说

代全部灰身长翅；再用子一代雄蝇与双隐性雌蝇黑身残翅交配，子二代有两种，灰身残翅及黑身长翅，比例各半，这一结果无法用 Mendel 自由结合规律解释。对此，Morgan 提出连锁假说，如图 7-3。而当子一代雌蝇与黑身残翅雄蝇交配时，则子二代有 4 种类型：灰身残翅、黑身长翅、黑身残翅、灰身长翅，分布比分别为：42%、42%、8%、8%；对此 Morgan 又提出同源染色体间的交换假说，从而总结出遗传学的第三规律：连锁和交换规律。后来，Morgan 及其学派又根据交换值估计连锁基因之间的距离、相对位置和线型排列，制订了基因图，创立了基因理论。由于 Morgan 学派在细胞水平上的卓越工作，为细胞的遗传学研究奠定了基础。后来 Muller 把 Morgan 开创的研究领域叫作“细胞遗传学”(cytogenetics)。

分子遗传学是从分子水平来研究遗传和变异，它是分子生物学的主要内容。现在已清楚遗

传和变异是由核酸的结构和性质所决定，因此分子遗传学的产生自然是和核酸的生物学意义的发现直接关联的。1928 年 Griffith 发现几种肺炎双球菌的突变株其多糖细胞壁的性质有所不同。这些差别可以用免疫方法鉴别出来,用肉眼也可以看到其菌落的形状不一样。菌株 S 形成光滑的菌落,而菌株 R 的细胞缺乏产生多糖的能力，故菌落表面起毛。菌株 S 导致小鼠患肺炎，而菌株 R 是不致病的。将活的R型肺炎菌株培养物和死的 S 型培养物注人鼠的体内，结果该鼠传染上了肺炎；说明在鼠体内繁殖的细菌培养物可以产生一种活的致病的 S 型肺炎球菌。由此得出一个结论,死的S菌株含有一种能将 R 型菌株转化成S型的因子。1944 年 Avery 等以提纯的 DNA 作细菌转化实验，在实验上证明了转化因子是 DNA。分子遗传学作为一门学科从此开始。Beadle 通过微生物遗传学研究提出一个基因一种酶的假说，即一种酶的缺陷是由于控制这种酶合成的一个特定基因突变的结果，这个假说的进一步发展，提出了一个基因一条肽链假说，更具有普遍意义。Beadle 的理论组成了分子遗传学的一个基本部分。1950 年 Chargaff 等通过色层分析表明在 DNA 中腺嘌呤(A)的摩尔数和胸腺嘧啶(T)的摩尔数相等，而鸟嘌呤(G)的摩尔数和胞嘧啶(C)的摩尔数相等。即 A=T,G=C。之后,Wilkins、Franklin、Waston 和 Crick 以 X-射线衍射分析方法详细研究了 DNA 的晶体结构,终于在 1953 年 Watson 和 Crick 根据对 X-射线衍射分析结果的解释并利用分子模型和双股 DNA 间的碱基互补配对原则,提出了 DNA 的双螺旋结构，DNA 复制的机制随即得到了初步的定性解释。Watson-Crick 双螺旋模型的提出建立了分子遗传学的基本理论,是分子遗传学发展史上的一个重大里程碑。Watson-Crick 当时对 DNA 复制的解释是半保留复制。Meselson 和 stahl 于 1958 年应用密度梯度离心技术,证实了 DNA 的半保留复制。Kornberg 在 1956 年发现了 DNA 聚合酶,并在以后研究中对 DNA 复制模型作了新的发展,提出了间断复制模型,以后为冈崎用实验证实。Weiss、Marmur 等对原核细胞的研究发现了 RNA 参与遗传信息转录,即 mRNA 的作用。Holly 提出了 Ala-tRNA 的模型。通过 Crick、Nirengberg、Khorana 及 Ochoa 的工作，阐明了遗传信息以三联体密码形式存在于核酸的碱基序列里,并通过碱基序列的转录、转译而指导特异性蛋白质的合成。细菌和病毒遗传学的发展，为分子遗传学的发展奠定了基础。五十年代开展的细菌遗传学研究,促进了大肠杆菌乳糖发酵中酶的诱导现象的深入研究,从而在 1961 年 Jocob 和 Monod 提出了乳糖操纵子模型。由于遗传和变异的不少基本问题在分子水平上得到了阐明，使我们有可能转向更复杂的问题,如细胞分化、胚胎发生、肿瘤形成；也使我们有可能进入到更深的层次——电子水平——量子遗传学。

早在 1935 年,Delbrück[6]就已指出量子理论和遗传学基本定律间有着密切的相似，染色体中遗传物质在漫长年代里的巨大稳定性仅仅说明它可能是一个处于稳定状态的巨大分子，那么突变也可能就对应于它经“量子跳跃”而变成的异构形式。量子力学创始人之一 Jordan 也主张“突变”是一种量子过程。Schrödinger[7]在他的那本“生命是什么”的著名著作中详细地讨论了 Delbrück的模型。他指出在染色体纤维中的基因可以适当地叫作一种非周期性晶体。并认为遗传密码应由有一定顺序的并在分子排列中具有稳定性的原子组合而成。他说：“实际上，在这样一种结构里,为了能形成几乎是无限多种可能的排列,其原子数目并不一定要非常大”。Schrödinger 认为,人的某些遗传特性可以在几百年里经过许多世代而被完整地保存下来。如果从哺乳

动物的起源算起，那么象胎生哺乳动物、体表被覆皮毛这些在遗传上固定的特征，已经在哺乳动物的系统发育史上经历了漫长的地质年代。而这些遗传特性的物质基础一直被保存在37°C左右的恒温条件下，好几千万年来分子热运动的影响居然没有破坏它的有序结构。因此外界因素如果要引起生物的遗传物质发生突变，必须经过一个量子化了的很高的能阶。早在1927年遗传学家 Müller 已发现X射线可以引起果蝇的突变。这一事实也许是后来 Schrödinger 在1944年提出上述看法的实验依据。物理学家 Gamow[8]在 Watson-Crick 提出双螺旋 DNA 结构一年之后，就致力于解释遗传密码问题。虽然 Gamow 当时的假说是错的，但他对这一重要问题所作的初步探讨却引起了物理学家进一步对分子生物学的关注。控制论的奠基人 Wiener[9]对遗传的理论问题也感兴趣，他提出了一个独特的想法："生物学的主要问题之一就是组成病毒的基因和产生癌的某种特殊物质是通过什么方式，从不具有特异性的氨基酸和核苷酸之类的混和物中繁殖自己。通常所作的解释是，这些物质的一个分子作为模板而起作用。这大都是一种形象的说法，只不过是对生命的基本现象的另一种描述方法。这是一个动力学过程，分子特异性的积极承担者可能是分子辐射的频率式样，这种辐射的一个重要部分可能是在红外电磁频率或更低的范围。一种病毒物质在某种情况下可能发射红外振荡，这种振荡具有促进从没有特异性的氨基酸和核苷酸原料中形成病毒分子的能力。很可能这种现象可以看作是频率的一种相互作用"。"许多储存信息的方法，共同具有一个重要的物理要素，这就是它们似乎都是高度量子简并性系统。换句话说，都是振动方式很多但频率相同的系统。……储存信息的许多物质，似乎都是具有大量内部共振的物质。量子简并性的出现，是由于这种物质具有某种由很小原因就能产生显著而稳定的结果的能力"。Wiener 还认为，高共振物质十分普遍地都有一种非常的储藏能量和信息的能力，新陈代谢和生殖作用的许多问题，都与量子简并性物质有关。1962年，Löwdin[3,10-13]根据隧道效应的概念，提出了 DNA 中氢键的质子穿透(proton tumeling)观点，用以解释 DNA 复制和突变。并首次提出了"量子遗传学"(quantum genetics)的名称。自六十年代以后，不少量子生物学的研究涉及到 DNA 双螺旋的结构稳定性、DNA 复制的量子化学计算、DNA 与蛋白质相互作用的计算、tRNA 的结构研究等等；虽然从总的讲来属于量子生物学范畴，但却都是从遗传学观点沿着 ⟲DNA —转录→ RNA —转译→ 蛋白质 这一中心法则展开的，我们可以把这些方面的工作归纳为遗传的量子化学研究或者量子遗传学。

遗传学从 Mendel 的表型水平、Morgan 的细胞染色体水平、Watson-Crick 的分子水平一直到目前处于萌芽状态的电子水平，经历了几个层次，这是与生物科学由描述性向精确性，由宏观到微观方向的发展相一致的。

目前，在国内外文献中"量子遗传学"这个名称还不多见，自 Löwdin 提出来之后响应者无几。也许是当时"量子遗传学"作为一门学科提出来还嫌过早。但是纵观二十多年来的量子生物学文献，有相当一部分内容可看作量子遗传学的。在量子生物学这一宽广的领域中划出一个量子遗传学的天地，这还是适时的可取的。本章下面几节将根据分子遗传学的框架，来归纳原子——电子水平上的遗传学研究文献。作者曾考虑过用"遗传学的量子理论"作为本章的标题，但 Löwdin

树起的“量子遗传学”目标明确，有更大的凝聚力，相对于研究者来讲也是一个更高的“激发态”，故取“量子遗传学”。但要指出，量子遗传学还没有一个完整的理论，仍是一个很不完备的体系。有人曾建议过从DNA大分子的电子激发态来着手研究量子遗传学，这是一个好主意[4]。作者在写本章时，正在美国作生物医学研究，利用了实验室的计算机终端(此终端与全 世界七 百多个大小数据库相连，并没有查到一本“量子遗传学”专著，因此本章 拟以 归纳为 主作 一个 初步 尝试。

7.1.2 分子遗传学概要[2,4,5,15-17]

7.1.2.1 基因及其复制

1）基因寓于DNA 下列实验可证明：

a．对每一物种而言，其一个细胞内之DNA含量是不变的，这说明DNA有作为遗传信息载体的条件；

b．紫外线可使基因产生突变，其谱线范围和核酸的吸收谱线范围相对应，并且还有其他一些已知可致突变的物理和化学作用亦可改变核酸；

c．细胞转化的因子是DNA；

d．大多数细菌病毒和许多动物病毒只含有DNA和蛋白质，而蛋白质只起到使 DNA 进入宿主细胞的作用。

DNA作为遗传的物质材料，必具有基因的基本性质，即它能精确的复制和具有 支配酶及其他蛋白质合成的能力。

2）DNA的半保留复制 DNA的复制是种半保留复制。DNA 复制时，其分子的 双链先解开成两条单链。然后以每一条原有的单链作为模板，各自互补地用细胞核内游离的核苷酸进行碱基配对，形成一条互补的新多聚核苷酸链。结果，一个亲代DNA 分子复制成两个结 构完全相同的子代DNA分子，每一个子代DNA分子均由一条原有的链和一条新链组成，即称 谓半保 留复制，如图 7-4 。

这一机制首先为以^{15}N标记的大肠杆菌实验所证实。即大肠杆菌先 置于^{15}N 培养基中繁殖，使之DNA完全为^{15}N所标记。然后于^{14}N培养基中待分裂一次后，将此进行过一次分裂的细胞取出一部分进行氯化铯密度梯度离心。结果发现并没有两种DNA，而是只有一种DNA，且位于重DNA(含^{15}N)带和正常DNA(含^{14}N)带中间；即表明此DNA的N是由等量^{14}N和^{15}N杂合而成，由此证实上述半保留复制机制。

3）原核细胞的DNA复制 目前对原核细胞的DNA 复制了解 得较清 楚。但 对真 核细胞DNA复制情况却远不及原核细胞这么详细。由于真核细胞DNA分子大，又与组蛋白 相结合，情况更复杂。此外，有些病毒的遗传物质是RNA，其复制是以RNA为模板。

原核细胞DNA复制过程主要是：

a．DNA解链：在DNA复制前，DNA解旋酶使DNA双螺旋解开；解旋蛋白结合到解开的DNA单链上，使DNA单链处于伸展状态。

b．引物形成：起始DNA的合成需要一种引物，一般是RNA片段。引物酶可识别 DNA的

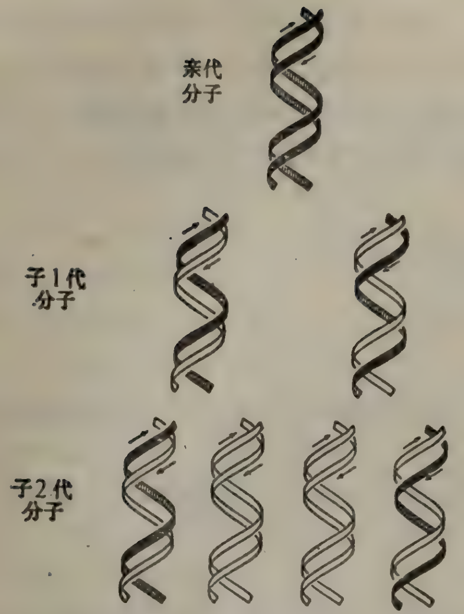


图 7-4 DNA的半保留复制

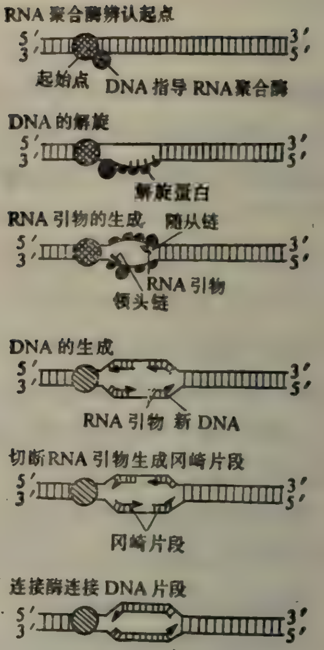


图 7-5 DNA间断复制过程

复制起点，诱发引物形成；再在 DNA 聚合酶催化下沿 5′→3′方向合成一条新的互补 DNA 链；DNA 复制完成后由核酸酶切去引物，最后由酶修补及连接成一完整DNA链。

c．间断复制：新链合成总沿 5′→3′方向，即核苷酸的聚合要从引物的 3′-羟基端开始，而DNA双链是反向平行的，因此理论上讲合成后为 3′→5′方向的新链的聚合需间断进行。实验结果是先生成冈崎片段，再由连接酶连接成完整的DNA新链，如图 7-5。

7.1.2.2 基因及其表达

1）蛋白质合成的总过程　　基因通过一系列复杂的生化过程最后表现为特定的性状，即为基因的表达。基因表达的二个最基本过程是转录和转译；基因表达的形式即蛋白质合成。蛋白质合成的主要过程是

DNA⟶RNA⟶蛋白质

如图 7-6，细胞核内的DNA的遗传信息首先转录到mRNA上；然后mRNA迁至核外，成为蛋白质合成的模板。核糖体结合在mRNA上完成蛋白质的合成。运载着特定氨基酸的tRNA进入核糖体内，tRNA上的反密码子和mRNA上的密码子以互补原理相识别；载着另一个特定氨基酸的tRNA再进入核糖体，二个tRNA上的氨基酸在酶的作用下以肽键相联；核糖体沿 mRNA 5′→3′方向逐步移动，如此逐个进行，而tRNA也逐个释放出来，完成蛋白质合成。

2）转录　　遗传信息从DNA转移到RNA称为转录。转录的过程是以DNA 分子中的一条单链为模板，按照碱基互补原则，合成一条互补链，即RNA，这种RNA接受DNA上的遗传信息，进入细胞质，参加转译过程，故称为信使RNA（mRNA）。

目前对原核细胞的转录了解得较深入。由依赖于DNA的RNA聚合酶催化的DNA 指导下的

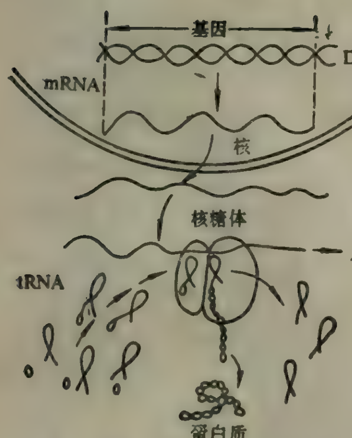


图 7-6　蛋白质合成过程

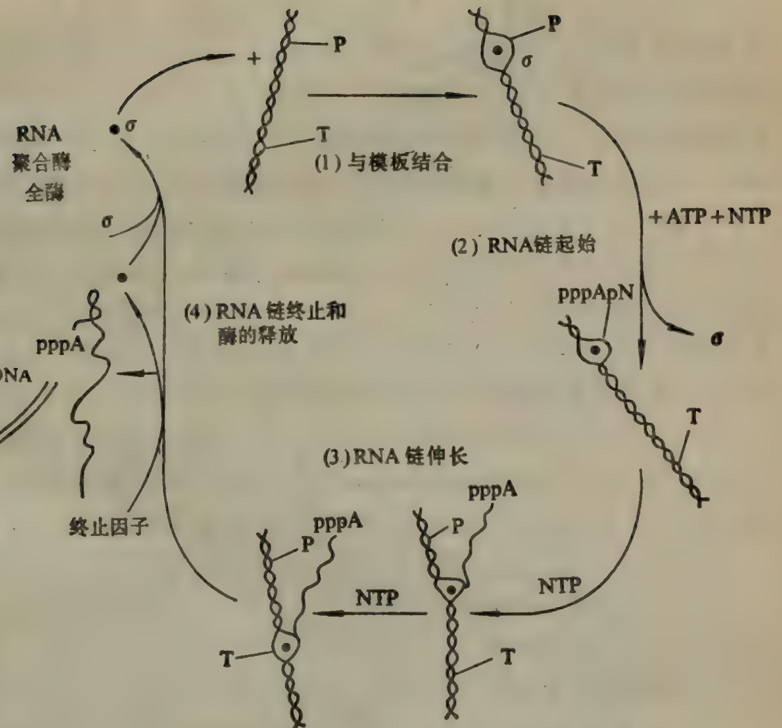


图 7-7　转录循环

RNA合成,即转录循环,由下列 4 个步骤组成：a．RNA聚合酶与模板结合；b．链的起始；c．链的延伸；d．链的终止和酶的释放。如图 7-7。在第一步,RNA聚合酶只有在两股 DNA 中某一链的某些部位才具有与模板结合的极高的亲合力。这些部位称为启动部位，只有在这些部位某一基因的转录才能开始;并且需要σ因子帮助去辨认启动区。第二步,ATP或GTP可作为起始核苷酸以生成一个二核苷酸四磷酸酯,当要进入第三步时σ因子即被释放。第三步中二核苷酸加上核糖核苷单磷酸酯使链伸长。第四步,当RNA聚合酶滑到终止信号时,它能识别基因终止信号,使转录停止,形成一条完整的mRNA。DNA模板和RNA转录本之间的相互作用是瞬间的,当RNA转录完成之后,DNA双螺旋结构全部复原,因此，转录与 DNA 复制不同，是全保留过程。

3）密码子　　基因的遗传信息通常是通过蛋白质的一定结构和生理生化特性表现出来的。蛋白质一般由 20 种氨基酸以不同的种类、数量、排列组合而成。蛋白质不能直接从核酸分子上以类似于转录的方式合成,因此需要从核酸“文字”翻译成蛋白质“文字”,故某些碱基的组合和某种氨基酸之间存在着对应关系。即mRNA的碱基序列中每三个相邻的碱基组成一个密码子;不同的碱基组合、构成不同的密码子,分别同一定的氨基酸相对应，决定有关氨基酸在蛋白质分子中的位置。密码子的特点是：a．它是mRNA上三个连续的碱基，又称三联体密码；b．密码子间无标点符号；c．某些氨基酸对应 2 个以上的密码子，即密码子有简并性；d．一个起始密码

子作为转译开始的信号,三个终止密码子作为转译完毕的信号。密码 子见表 7-1。在不 少情况下,简并性表现在一个氨基酸对应于几个具不同反密码子的 tRNA上。在另一 些情 况下，一个tRNA可以识别几个密码子,如表 7-2,如$tRNA^{Gly III}$可以翻译GGU和 GGC，但不能翻译其他两种甘氨酸密码子;而$tRNA^{Gly II}$却可翻译另两种密码子,情况正相反。另一方面,$tRNA^{Gly I}_{GGG}$只能翻译GGG。由于前两种甘氨酸的tRNA存在,似乎$tRNA^{Gly I}$无存在的必要。从表 7-2 也 可以看出当一种反密码子可以识别一个以上的密码子时,这些密码子的差别仅在 3′- 端的位置上;这一特征可用摆动假说(wobble hypothesis)解释。即密码子的第三个碱基可以形成非互补的碱基对,而这时其构象接近于标准互补对A-U,G-C。变位假说预示了一些碱基对,如表 7-3。一般说来,变位假说预示的碱基对已证明是正确的,但有一例外，即表 7-2 中的 $tRNA^{Gly III}$ 中的 反密码子 $^{3'}CCA^{5'}$是由甘氨酸密码子$^{5'}GGC^{3'}$所识别。

表 7-1　遗传密码子

第一字母	第二字母 U	第二字母 C	第二字母 A	第二字母 G	第三字母
U	UUU } 苯丙氨酸 UUC UUA } 亮氨酸 UUG	UCU } 丝氨酸 UCC UCA UCG	UAU } 酪氨酸 UAC UAA 终止 UAG 终止	UGU } 半胱氨酸 UGC UGA 终止 UGG 色氨酸	U C A G
C	CUU } 亮氨酸 CUC CUA CUG	CCU } 脯氨酸 CCC CCA CCG	CAU } 组氨酸 CAC CAA } 谷氨酰胺 CAG	CGU } 精氨酸 CGC CGA CGG	U G A G
A	AUU } 异亮氨酸 AUC AUA AUG { 甲硫氨酸 起始	ACU } 苏氨酸 ACC ACA ACG	AAU } 天冬酰胺 AAC AAA } 赖氨酸 AAG	AGU } 丝氨酸 AGC AGA } 精氨酸 AGG	U C A G
G	GUU } 缬氨酸 GUC GUA GUG	GCU } 丙氨酸 GCC GCA GCG	GAU } 天冬氨酸 GAC GAA } 谷氨酸 GAG	GGU } 甘氨酸 GGC GGA GGG	U C A G

表 7-2　甘氨酸tRNA的多重性

甘氨酸 tRNA	tRNA 中的反密码子	tRNA 识别的甘氨酸密码子
$tRNA^{Gly I}_{GGG}$	$^{3'}CCC^{5'}$	$GGG^{3'}$
$tRNA^{Gly II}_{GGA/G}$	$^{3'}CCU^{5'}$	$^{5'}GGA^{3'}$ $^{5'}GGG^{3'}$
$tRNA^{Gly III}_{GGU/C}$	$^{3'}CCA^{5'}$	$^{5'}GGU^{3'}$ $^{5'}GGC^{3'}$

遗传密码可能是具有普遍性的。这结论主要来自:(1) 通过分析不同的 蛋白 质如大肠杆菌色氨酸合成酶突变后的 α 链和人血红蛋白突变后的 α 链的氨基酸顺序之变化;(2) 测定 各种物种的mRNA的碱基对顺序。从研究同源蛋白质得到的结果，表明即使在关系很远的物种之间其

表 7-3 符合变位假说的碱基对

反密码子上第三位的碱基	密码子上第三位的碱基
G	U or C
C	G
A	U
U	A or G
I	A, U, or C

同源蛋白质的氨基酸顺序有相当大部分是保留下来的，这亦支持遗传密码普适性的观点，例如在动物、植物和真菌等各种物种的真核细胞色素c之间，约有 1/3 的残基顺序是共同的。

已发现在线粒体基因组中，遗传密码普适性有了例外。对人类线粒体 DNA的序列分析，表明至少有 4 种密码差别：AUA通常是异亮氨酸的密码子，但这里却对应了蛋氨酸，UGA 通常是终止信号，但这里对应于色氨酸；AGA和AGG通常是对应于精氨酸，但这里却为终止信号，而AUA或AUU则代替了AUG，成了起始密码子。这些密码变化的基础还不了解，然而遗传密码变化的一个结果是使线粒体的遗传隔离。当一个RNA含有上述任何一种密码子时，在线粒体膜内和线粒体膜外被转译时，很明显将会对应于不同的氨基酸。目前关于遗传密码的起源并不清楚。在tRNA反密码子或mRNA密码子和其对应的氨基酸之间尚不知有化学上的逻辑联系。

4）转译　基因的遗传信息从mRNA的核苷酸顺序翻译成相应的多肽链氨基酸顺序，这个过程为转译。目前，原核细胞转译机制比较清楚。模板mRNA碱基顺序译成蛋白质的氨基酸顺序主要有两步：(1)氨基酸的激活和转运；(2)mRNA“解读”与蛋白质合成。氨基酸分子的活化即它的羧基与ATP反应，形成氨酰腺苷酸，再和相应的tRNA(转移tRNA)3′端的腺嘌呤残基结合，形成氨酰-tRNA。原核生物中在起始转译时起决定作用的是甲酰甲硫氨酰-tRNA。

“解读”即是tRNA上的反密码子和mRNA上的密码子互相识别。蛋白质合成是在核糖体上进行的。大肠杆菌的核糖体是 70 *S*（*S*为沉降系数）的，但在Mg^{2+}存在下，则分离为二个亚单位，即 30 *S*和 50 *S*。30 *S*、50 *S* 的亚单位先后结合到 mRNA 上，形成 70 *S* 的核糖体。70 *S* 核糖体有两个专门结合 tRNA 的部位，A位氨基酰部位是结合氨基酰 tRNA 的反密码子和该位上的 mRNA 的密码子相配合的部位，P 位肽酰部位是只能接受带肽键的 tRNA 的部位，即接受从 A 位转移的 tRNA 分子。第一步是起始的甲酰甲硫氨酰-tRNA 的反密码子 UAC 同 mRNA 中的起始密码子 AUG 互补；再是第二个氨酰-tRNA 结合到核糖体的 A 位上，核糖体 P 位上的氨基酸的羧基和 A 部位上的氨基酰-tRNA 的氨基形成肽键。这时 P 位上脱掉氨基酸的 tRNA 便从核糖体上释放出来，进行下一次循环。核糖体沿 mRNA 5′→3′方向移动一个密码子。使原来在 A 位上的带有肽键的 tRNA 移至 P 位，而 mRNA 的密码子处在 A 位上以接受下一个氨酰-tRNA，如此多肽键逐渐延长，直到终止密码子，释放出肽键，合成一种特定的蛋白质，如图 7-8 。

7.1.2.3　基因表达的调控

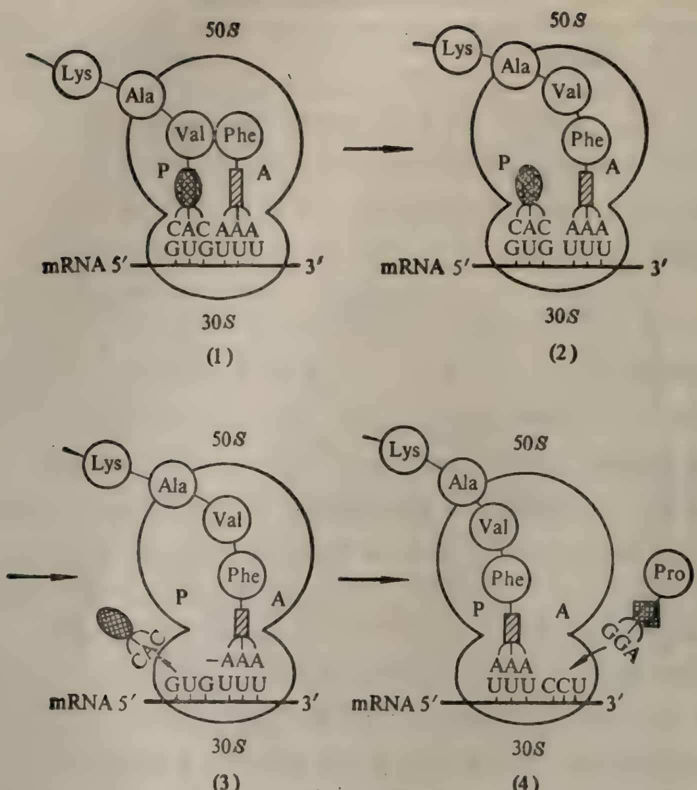


图 7-8 多肽链在核糖体上生长

1)基因表达控制的概念 基因表达的调控可以在从基因到多肽产物生成的途径中之任一环节发生。一般可分为转录控制、转译控制和反馈抑制。到目前为止,对细菌代谢的遗传调控研究得较深入,而对于真核生物的基因调控问题只有一些假说,都还缺乏足够的实验证据。已知在转录水平上的调控系统有负调节和正调节。负调节是某一细胞成分的存在使某种细胞功能停止,而当这一成分消失或失活,则使这一功能恢复。属于负调节的例了是乳糖操纵子模型和色氨酸操纵子模型,这里酶的停止合成是由于调节基因的产物——阻遏蛋白作用的结果。当阻遏蛋白和诱导物相结合时,它就失去活性,不能再和操纵基因相结合,酶合成便能进行。正控制是某一细胞成分的存在使某种细胞功能实现,而这一成份的消失或失活使这一功能停止,这里酶的合成必需有一种调节基因的产物——降解物激活蛋白(CAP)的存在。转录或转译控制的对象都是酶的合成。反馈抑制是控制酶的活性,其作用是合成代谢的最终产物对于合成途径中的第一个酶的可逆性的抑制作用。转录控制或转译控制都涉及调节基因所指导产生蛋白质的作用。反馈抑制因是代谢最终产物直接和酶结合而显示其抑制作用,所以不存在调节基因,而与编码第一个酶的结构基因直接有关。

2)乳糖操纵子模型 大肠杆菌中的乳糖代谢时的基因调控研究得最清楚,可用乳糖操纵子模型来解释。乳糖操纵子(operon)是由启动基因 P、操纵基因 O 和三个结构基因 Z、Y 和 A 所

构成的功能单位，如图 7-9。当操纵基因 O 同调节基因 i 的产物阻遏蛋白相结合时，转录便不能进行；当阻遏蛋白同乳糖等诱导物相结合时，脱离操纵基因，转录可以进行。

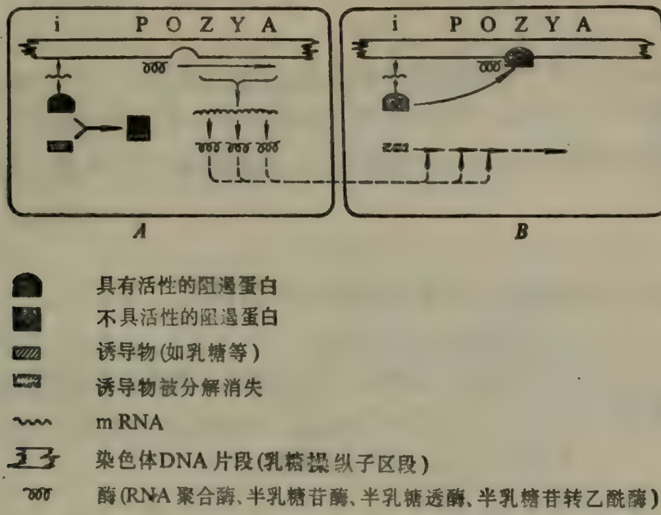


图 7-9 乳糖操纵子控制基因表达的示意图

3)基因突变　DNA 中碱基顺序的变化将会引起永久的遗传性的改变，即基因突变。染色体组中的 DNA 的任何变化将会反映在 mRNA 上以致影响合成的蛋白质，或者是反映在 rRNA 和各种 tRNA 上。点突变是指一个碱基或一个碱基对的变化，这样经过 mRNA 的转录，最终使肽链顺序中的一个氨基酸被另一个取代。点突变是最常见的基因变化，它是由 DNA 复制或修复时发生的错误引起的。它可由化学或物理因子作用而发生。

以上简单介绍了原核细胞的分子遗传学概念。其核细胞由于存在核膜及多种亚细胞器，其转录、转译、复制机制也更复杂。目前了解比较仔细的是原核细胞的情况，量子遗传学目前的任务主要是试图来解释已经比较清楚的原核细胞的遗传学问题。下面我们来讨论量子遗传学的主要内容，即亚分子水平(submolecular level)或原子和电子水平的信息流。

7.2 原子和电子水平的信息流

7.2.1 基因复制和突变的微观机制

7.2.1.1 遗传密码的基本结构

我们知道 DNA 只有 4 种碱基：A、T、G、C，现在要问为什么 DNA 里只有 4 种碱基？理由是：a. 遗传信息寓于 DNA 的碱基顺序，即碱基顺序的不同排列组成了遗传密码，也即每种碱基是个编码符号；当在一定的约束条件下，对由符号的线性排列构成的信息量求极大时，可导得最优符号数是 3。b. 碱基互补复制要求符号数为偶数。因此满足这两个条件的符号数只可能是

4^N。这可从信息论来给予证明[18]。

在讨论生物的结构与功能时，有一个很重要的观点：在进化过程中，机体只有选择其最大效率的结构时才能优于其竞争对象而得以生存；即能存在的结构其相对效率将是最高的。在下面的证明中将要用到这一观点。

假定一信息码(code)由几个线性排列的单位构成，而这码的每一单位的符号又有 N 种可能的等同的选择。再假定这一编码的"成本"正比于码的长度 n 和符号的数目 N，即"成本"正比于 nN，把 nN 视作"符号信息贮存池(pool)"的面积。由排列组合可知，信息码的可能排列数 D 为下式决定：

$$D=(N)^n \tag{7.1}$$

这时信息码的信息量 I，根据信息论，可作如下计算：

$$I=\log_2 D=n\log_2 N \tag{7.2}$$

在约束条件为 $nN=\text{const}$ 时，对函数 I 求极大，即可得信息量最大时的符号种数 N；因此，对函数 I 运用Lagrange 乘子，

$$F(N,n,\lambda)=N^n+\lambda(nN-C) \tag{7.3}$$

解为：$InN_0=1$ 或

$$N_0=e=2.72$$

这里 N_0是在编码中使用的符号的最佳数目，考虑到符号数目只可能是整数，因此 $N_0=3$。到这里，我们仅考虑了 DNA 的信息贮存的作用方面；还需考虑 DNA 的信息复制的作用方面。

根据 Watson-Crick 碱基对互补原则，碱基数目一定要是偶数的，即成对的。每一碱基需要有一个与它互补的碱基相匹配。这就是说符号数目必须为偶数，结合上述 $N_0=3$ 的条件，可知三种碱基只能形成一对半互补对，还缺一种碱基，因此 N 只有等于 4，即 DNA 中碱基数目为 4 时，DNA 的功能将有最大的"效率"，既可发挥信息贮存作用，又可发挥信息复制作用。

现在我们来讨论遗传密码的基本结构，前面已提到，细菌转化实验确认了遗传信息寓于 DNA，通过色层分析及 X-射线衍射分析建立了 DNA 的双螺旋结构及碱基互补原则：

A-T　　G-C

本节中要讨论的遗传密码指的是 DNA 中的碱基及其排列方式：

Watson-Crick 的碱基对互补原则基于下面二个条件：(1)立体条件：在 DNA 双螺旋中，嘧啶只有和嘌呤相连；(2)氢键条件：碱基对之间必须至少有二个处于合适位置的质子，以形成至少二个氢键。

根据条件(1)、(2)，可以确定，在 DNA 的四种碱基中只有 A-T、G-C 组合的可能。但是，Delbrück[6]和 Schrödinger[7]早已指出，如果在某种大分子里存在遗传密码，则和突变相对应的这种遗传密码里的不连续变化可能决定于这种大分子是否存在互变异构。在 DNA 中的碱基，有简单的互变异构现象（见 4.2.2)。

我们知道，互变异构的变化，是因为质子在该碱基中所处位置的不同；而氢键的本质即是一个质子共享二个孤对电子，即：:H⋯: 或 :⋯H:。因此，我们可以把注意力集中在碱基的质子、电子对上，采用 Lewis 记号，这样遗传密码的基本结构就以质子-电子对表征出来了：



而其相应的互变异构体编码即为



如果我们再用互补原则的二个条件来衡量这四种碱基及其四种异构体中所可能的碱基对结构，则发现也可以有下列配对方式：

A*-C,　　A-C*,　　G*-T,　　G-T*

这就意味着在 DNA 复制过程中，如果由于某种原因，在模板 DNA 分子或营养物质中出现互变异构形式的碱基，则在拷贝过程中会引进一个错误。这样，使原来应该是 G-C 对的地方变成了 A-T 对，或原来应该是 A-T 对的地方却变成了 G-C 对；这是因为，一个复制错误，会一步步地经生物"放大"，最终很快会变得可以识别，这就是 Watson-Crick 提出的突变机制，图 7-10：

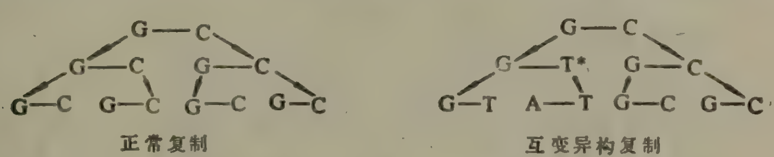


图 7-10　Watson-Crick 突变机制

7.2.1.2　DNA Y 型复制的复制平面模型

首先要说明，在这里我们讨论的是一般双螺旋的 DNA 的复制。有些噬菌体，如 ϕX 174、M_{13}系单股 DNA，另有其复制机制。

原则上来讲，DNA 复制的三种机制：保留复制、半保留复制及分散复制都是可能的[19]。当初 Watson-Crick 提出的复制模型过于简单；事实上，DNA 复制涉及旋绕-解旋机制，这是个十分复杂的拓扑问题。由于从理论上讲，DNA 复制机制有许多存在的可能性，因此只有用实验去确认何种机制真正存在。1958 年，Meselson 和 Stahl[20]的实验确认了复制过程中遗传材料是半保留的；后来 Taylor[21]又从蚕豆染色体复制的研究中进一步证实了此结果，但这在具体细节上还要作许多解释。1955 年 Delbrück[22]，后来又和 Stent[19] 提出了 DNA 复制的 Y 模型，然而 Kornberg[23,24]的离体实验表明，核苷酸只能从聚核苷酸链的自由糖端加上去，又因 DNA 的两条链是反向的，因此曾一度认为 Y 型模型的复制方式不可能存在。原来 Cairns[25]用氚标记胸苷并用放射自显影术作了大肠杆菌复制实验，令人信服地证明复制机制是Y 型的并涉及旋绕-解旋过程。所以当时已知线形的双螺旋 DNA 是 Y 型复制。遇到的问题是根据 Y 模型，DNA 解旋之后，单股 DNA 由于糖-磷酸酯骨架有许多单键，将会有很大的自由度。Watson-Crick 结构碱基互补的主体条件，即嘌呤针对嘧啶在双螺旋内部是必要的；但在单股 DNA 情况下，这样的立体条件是否仍属必要，就很难讲了，这是否意味着碱基配对时只有氢键条件起作用？事实上，

碱基对自发突变的机率非常小，每代只有 10^{-8} 至 10^{-11}。那么可以设想，在 DNA Y 型复制时，一定有着某种机制维持着 Watson-Crick 碱基对互补原则。

基于这一考虑，Löwdin[3]提出了 DNA 复制的复制平面模型，即 DNA 复制选择原理。Löwdin 基于 Rich[26] 的多股螺旋存在的观点，设想有一个 DNA四股螺旋，但这四股螺旋只在某一碱基平面上可以存在，这平面位于 Y 模型的交叉点上。这平面就称作“复制平面”。如图 7-11，h_1、h_2、h_3、h_4 即四股螺旋。Löwdin 根据对在这一平面上的氢键成键条件的分析，得到了 DNA 合成时的选择原则。下面，我们详细介绍这个模型。

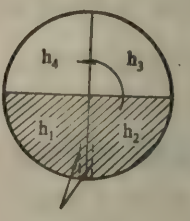
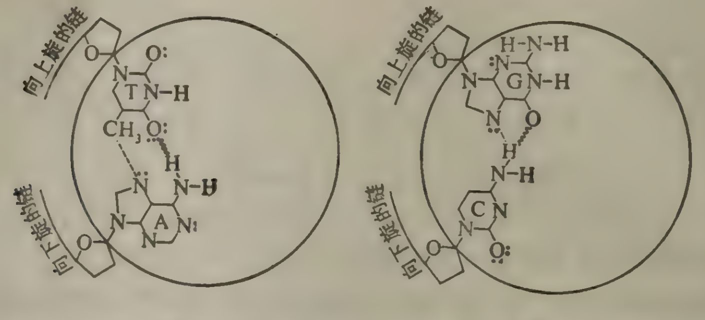


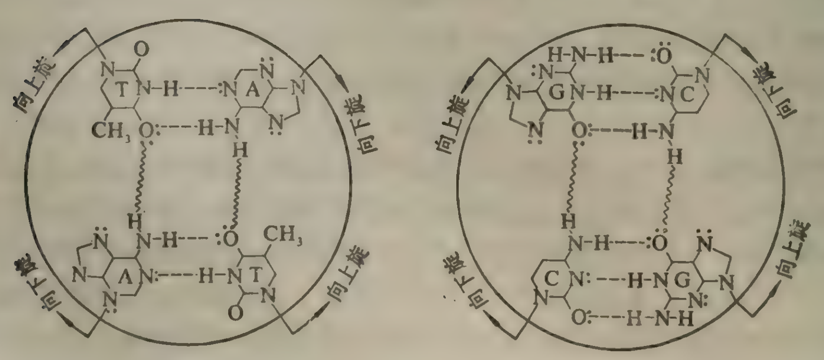
图7-11　复制平面，一股 DNA 绕轴旋转 180°使质子码打开

一般的双螺旋，每一碱基相对于螺旋轴的角度大约是 90°；当有四股螺旋时，而其糖磷酸骨架之间又没有任何的密堆积，这样四股螺旋所占的总空间要比双螺旋大得多。在复制开始前母



A　　　　B

图7-12　母体 DNA 的股间氢键打开使 h_2 位置中的一股 DNA 转至 h_4位置，使质子-电子对码处于开启状态，这时有一个暂时氢键形成，图中波纹线即是



A　　　　B

图7-13　新的互补碱基配对之后的复制平面，这时形成二对暂时氢键（波纹线所示），并可起“纯化”作用

体 DNA 的二股螺旋处于 h_1、h_2的位置。现在先设想一种最简单的情况：使这二股螺旋碱基间的所有氢键同时断开，并使整个 h_2 螺旋绕轴自由旋转至 h_4的位置，再与 h_1上的碱基形成一个新氢键。这样，相应的质子-电子对码处于开启状态，并且这二股糖-磷酸骨架移动的位置要比原来的更向外面一些，使得核苷酸容易进入复制位置。同时，质子码仍被螺旋屏蔽，不至受外界环境的干扰，如图 7-12。在这个最简单的模型里，是以整个母体双螺旋作为模板的，因每一平面只有一个嘌呤及一个嘧啶，因此互补原则所要求的主体条件仍然有效；这样，每一碱基只接受一个符合氢键条件及主体条件的配对碱基，如此就形成一个四碱基的平面，如图 7-13。

现在，基于上面这一最简单的模型，再考虑一个进一步的模型，即只采纳"复制平面"的概念。设想此平面从母体双螺旋的一端一步步地移至另一端，从而使二股螺旋逐步解开；或者也可以看作复制平面基本静止而母体双螺旋作螺旋运动，这可能与 Y 模型更符合些。这里要指出，处于复制平面上的二股分开的螺旋的位置（图 7-11 的h_1、h_2 位置）为一暂时性的氢键所稳定，并且嘌呤 N 7 位上的孤对电子和 CH_3 基或 NH_2 基上的质子间也可形成氢键，这一氢键对于稳定处于 h_1、h_2 位置的二股螺旋也起作用。在这个进一步的模型里，母体螺旋上的密码对每次只在一点打开，即只在复制平面上打开，而其余大部分母体双螺旋仍处于原来的正常位置，只是在复制平面邻近的区域里的那些母体结构处于一个中间状态：其糖-磷酸骨架稍作变形从而适应螺旋的解旋。在这一过程中，由于分解和合成同时进行，故不需要很大的起始能量；每一复制平面上，当互补碱基配对时，有 5～7 个新键形成，此时释放出来的能量可用于打开母体螺旋上的 2～3 个氢键及在前一复制平面上的二个暂时氢键，如图 7-14。复制平面不仅提供了 Watson-Crick 结构互补性所要求的严格立体条件，并且还对用于合成的核苷酸材料起纯化的作用，以避免复制时引入错误。这可以这样来看，即在复制平面中部有一个特征的平行四边形，二个氨基，二个酮基分别处于对角。此时，这些基团是双重配对的，这种双重配对对于防止一个在环境里的互变异构体碱基引入新合成的 DNA，有重要作用。现以一个 A-T 复制平面为例，如图 7-15，只考虑关键基团，如只根据互补要求的二个条件，有三种碱基配对可能性存在，但是$\binom{\mathrm{T\ A}}{\mathrm{A\ C*}}$情况则二个质子相遇，在$\binom{\mathrm{T\ G*}}{\mathrm{A\ T}}$情况则二个孤对电子相遇，它们之间都有很大斥力，结构不可能稳定，只有$\binom{\mathrm{T\ A}}{\mathrm{A\ T}}$情况，可以形成额外的临时的氢键，而使结构稳定。如此就限制了一个互变异构碱基引入 DNA。这就是双重配对的"纯化"作用。但是，这种复制平面机制并没有排斥一对互变异构体引入 DNA 的可能性，不过这种机率很小。

Löwdin 的复制平面模型（即 DNA 复制选择原理），解释了为何存在 Y 型 DNA 复制时仍能维持其精确性。Löwdin 的模型是 60 年代中期提出来的，这一模型并没有去解决 Y 复制时的核苷酸只能从聚核苷酸链的自由糖端加上去的矛盾。当时认为，Kornberg 的实验结果是在离体条件下得到的，可能在体内有不同合成机制，从而回避了这一矛盾。现在我们知道，新链合成总是从 5′→3′ 方向进行的，冈崎[27]等提出 DNA 片段合成的间断复制机制，并用实验给以证实，在分子水平上解决了这一矛盾。但是间断复制在电子水平的机制尚不清楚，这个问题向量子遗传学提出了挑战。

7.2.1.3 遗传密码的质子-电子对模型

Löwdin[3]提出了遗传密码的质子-电子对模型，这是用量子力学方法来处理 DNA 复制和

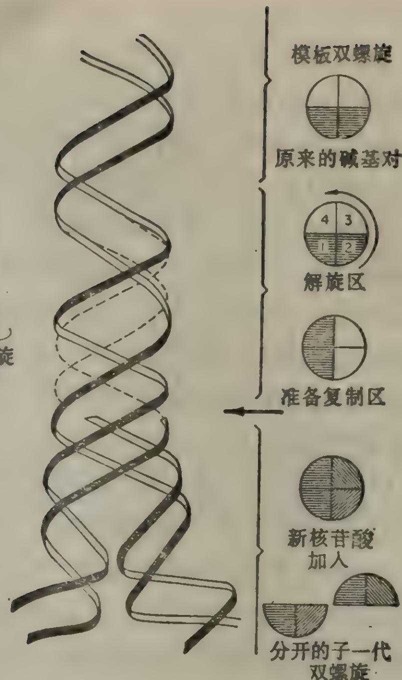


图 7-14　符合 Y 型 DNA 复制的机制

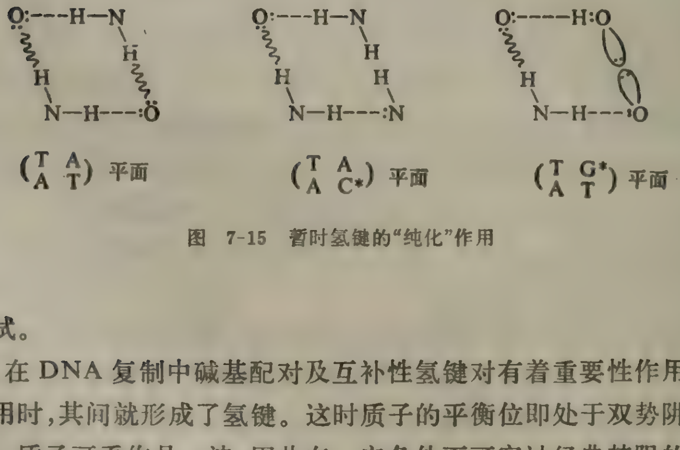


图　7-15　暂时氢键的"纯化"作用

突变机制的首次尝试。

我们已经知道，在 DNA 复制中碱基配对及互补性氢键对有着重要性作用。当一个质子与二个电子对相互作用时，其间就形成了氢键。这时质子的平衡位即处于双势阱的最小势阱中，图 7-16（参见 3.2.2）。质子可看作是一波，因此在一定条件下可穿过经典禁阻的势垒，即有遂道效应。从遗传密码的质子-电子对结构可知，质子在一对互补碱基间的转移，将形成相应的互变异构体，从而影响 DNA复制的精确性，即产生突变，这就是 Löwdin 模型的基本思想。这种质子转移的可能性决定于其隧道效应的穿透机率。在遗传密码里的质子，是与环境隔开的，但热运动仍可通过化学键传到质子上，另外质子在热运动的影响下也有可能越过势垒而发生转移，因此在计算质子在碱基间的转移时，也要考虑热运动。不过可以把这二种因素分成二步来考虑，先考虑隧道

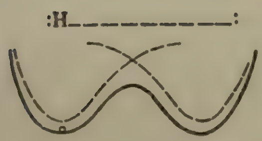


图 7-16 氢键的双势阱

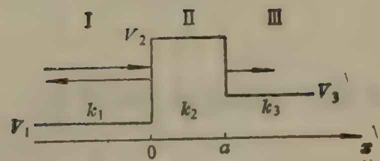


图 7-17 方势垒

效应的因素，再将热运动的因素加进去，然后计算总转移。在这里先简单介绍隧道效应。

考虑厚度为 a 的方势垒，如图 7-17。这是一个一维体系，相应的 Schrödinger 方程为

$$-\frac{h^2}{8\pi^2 m}\frac{\partial^2\psi}{\partial x^2}+V\psi=E\psi \tag{7.4}$$

令

$$p=\sqrt{2m(E-V)},$$

$$K=\frac{p}{\hbar}=\frac{2\pi}{h}\sqrt{2m(E-V)} \tag{7.5}$$

则式(7.4)可写作

$$\partial^2\psi/\partial x^2=-K^2\psi \tag{7.6}$$

现考虑 $V_3<E<V_2$ 的情况。当处于区域 I 的入射波射至势垒时，部份被反射，部份透射至区域 III。透射波强度和入射波强度之比即为势垒的透射系数。波函数在区域I、II、III的解分别是：

$$\begin{aligned}
&(\text{I})\quad \psi=\exp(ik_1x)+b\exp(-ik_1x); && k_1=\frac{2\pi}{h}\sqrt{2m(E-V_1)}\\
&(\text{II})\quad \psi=\alpha\exp(+|k_2|x)+\beta\exp(-|k_2|x); && k_2=\frac{2\pi}{h}\sqrt{2m(V_2-E)}\\
&(\text{III})\quad \psi=c\exp(+ik_3x); && k_3=\frac{2\pi}{h}\sqrt{2m(E-V_3)}
\end{aligned} \tag{7.7}$$

因区域 I 的入射波系数选为 1 （即(I)中前项为入射波项，后项为反射波项），故透射系数为 $|c|^2$。波函数在边界 $x=0, x=a$ 处是连续的。引入量 K

$$K=a|k_2|=\frac{2\pi a}{h}\sqrt{2m(V_2-E)} \tag{7.8}$$

消去 b、α 及 β 以后，可得如下关系

$$c^{-1}=\frac{\exp(ik_3a)}{4i}\left\{\left(\frac{k_3}{|k_2|}-\frac{|k_2|}{k_1}\right)(e^K-e^{-K})+i\left(1+\frac{k_3}{k_1}\right)(e^K+e^{-K})\right\} \tag{7.9}$$

因此透射系数的倒数为

$$|c|^{-2}=\frac{1}{4}\left(1+\frac{k_3}{k_1}\right)^2+\frac{1}{4}\frac{(k_1^2+|k_2|^2)(k_3^2+|k_2|^2)}{k_1^2|k_2|^2}\sinh^2K \tag{7.10}$$

此式表示透射随量 K 作指数衰减。上述公式是从方势垒导得的，当势垒为任意形状时，可采用 WKB 方法解出透射系数 g。

$$g=|c|^2\approx e^{-2K} \tag{7.11}$$

这里

$$K=\frac{2\pi}{h}\int_0^a |p|\,dx \tag{7.12}$$

若势垒顶部近似为抛物线形，则可导得下式

$$K=\frac{\pi}{4}\,\frac{a}{h}\sqrt{2m(V_{max}-E)} \tag{7.13}$$

以下讨论均假定势垒顶部为抛物线形，即用上式(7.13)来计算K。

先看对称双势阱，如图 7-18，并假定最初质子在左边阱中，并具有能量E，势垒高度为$V_0=V_2-V_3$，引入系数κ，则$E=V_2-\kappa V_0$。隧道距离$a=a(\kappa)$，很明显a是一和势垒形状有关的参数κ的函数。作为初步近似，用抛物线公式$a=a_0\sqrt{\kappa}$代入式(7.13)，得：

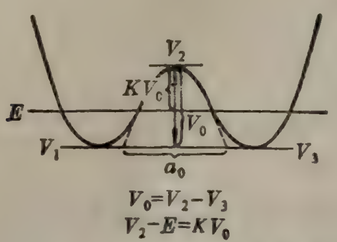


图 7-18　对称双势阱

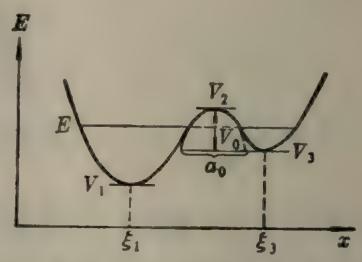


图 7-19　不对称双势阱

$$K=K_0\kappa,\qquad K_0=\frac{\pi}{4}\,\frac{a_0}{h}\sqrt{2mV_0} \tag{7.14}$$

从而透射系数为

$$g\approx 10^{-K'_0\kappa},\qquad K'_0=2K_0\log_{10}e \tag{7.15}$$

即可算得

$$g=10^{-14.992\kappa a_0 V_0^{1/2}} \tag{7.16}$$

这式中a_0用Å单位，V_0用eV单位。

现再来看非对称双势阱的情况，如图 7-19。如果$V_1<E<V_3$，则波函数在区域 II 及 III 内作指数衰减，不会有隧道穿透；如$V_3<E<V_2$，则可由式(7.16)计算透射系数，这里$V_0=V_2-V_3$，而a_0是最小穿透能级$E=V_3$时的势垒宽度；如果$E>V_2$，则质子越过势垒。

现在我们主要来讨论有穿透的情况。令质子在左势阱的势壁间来回振动之特征频率为ν_1；质子每一次碰到势垒时，其透射系数为g，则单位时间内质子穿透到右势阱的机率是

$$\nu_1 g=\nu_1 10^{-14.992\kappa a_0 V_0^{1/2}} \tag{7.17}$$

如果势阱底近似于抛物线形，则至少在最低穿透能级可对ν_1作一粗略的估计。Hamilton 量可写作：

$$\mathscr{H}=\frac{p^2}{2m}+\frac{m}{2}(2\pi\nu_1)^2(x-\xi_1)^2 \tag{7.18}$$

若 $2b_1$ 是质子在基态时在势壁间来回振动的二个经典回折点间的距离，则

$$E_0=\frac{m}{2}(2\pi\nu_1)^2b_1^2=\frac{h\nu_1}{2} \tag{7.19}$$

即

$$\nu_1=\frac{h}{4\pi^2mb_1^2} \tag{7.20}$$

若 b_1 用 Å 作单位，则

$$\nu_1=\frac{1.0013}{b_1^2}\times10^{12}\text{sec}^{-1} \tag{7.21}$$

我们可以初步估计一下 ν_1 的值，一般 $b_1<1$ Å，假定 $b_1\approx0.3$ Å 则 $\nu_1\approx10^{13}$ sec^{-1}，这个数量级是合理的。而能级 $E=(n+1/2)h\nu_1$ 的间隔为

$$h\nu_1=\frac{4.1408}{b_1^2}\times10^{-2}\text{eV} \tag{7.22}$$

其数量级在 0.04 eV。最后，可以导得质子在单位时间内的透射机率：

$$\nu_1g=\frac{1.0013}{b_1^2}\times10^{12-14.992\kappa a_0V_0^{1/2}} \tag{7.23}$$

此式中的 a_0、b_1 要用 Å，V_0 要用 eV 单位。还应指出，当 $\kappa\approx0$ 或 $\kappa\approx1$ 时，式 (7.23) 偏差较大。

现在，我们进一步来考虑热运动对质子转移的影响。氢键内的热运动、热平衡问题非常复杂，故这里只能粗略地假定质子在氢键里的热平衡可用 Boltzmann 分布来描述。

$$e^{-E/\kappa T}$$

并假设质子最初处于左势阱中，并已达到热平衡，如图 7-20。

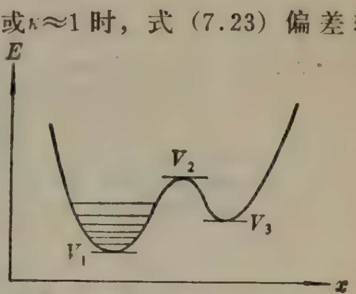


图 7-20 热平衡时质子在势阱里有了Boltzmann 分布

先来考虑势阱里和势垒上面的质子的 Boltzmann 分布，将 Boltzmann 因子化作以 10 为底的幂，以便估计数量级：

$$\exp[-(V_2-V_1)/kT]=10^{-K''(V_2-V_1)} \tag{7.24}$$

这里 $K''=\log_{10}e/kT=5040/T$，在正常人体温度下，T 310 K 则 $K''=16.26$，而在冰点和沸点时的 K'' 分别为 18.46 和 13.51。再计算一下差一个能级时的质子 Boltzmann 分布，能级间隔取 0.04 eV (前面已计算过)，T 取 310 K，这时 Boltzmann 因子是 $10^{-0.6504}\approx1/4$，这意味着在平衡时，质子的分布机率会大部分集中在势阱的底部 (计算时 $b_1\approx0.3$ Å)。当能级 $n=10$ 与能级 $n=0$ 相比时，其分布机率之比将是 10^{-6}。若势垒高度 $V_2-V_1=1$eV，则相应的 Boltzmann 因子将是 10^{-16}，这表明在势垒以上质子分布的可能性极小。因为势垒变化 0.1 eV，相应的Boltzmann 因子却要变化 $10^{-1.6}$，所以能量计算要非常精确。

现在来看考虑了热运动之后，质子的转移。原则上讲，计算时要考虑穿透的和因热运动而越过势垒的二个因素。这里仅计算穿透情况，若粗略地计算，可将 Boltzmann 因子乘以一个质子

在每秒内之透射几率 $\nu_1 g$，则得考虑了 Boltzmann 分布后的穿透转移：

$$Q(\kappa)=\nu_1 g\exp[-(E-V_1)/kT] \tag{7.25}$$

对于非对称双势阱的情况，来考虑一下最低穿透能级 $E=V_3$ 时的转移，这时 Boltzmann 因子为

$$\exp[-(V_3-V_1)/kT]$$

对于 $T=310$ K，Boltzmann 因子是

$$B=10^{-16.26(V_3-V_1)} \tag{7.26}$$

因此考虑了Boltzmann 分布后的质子穿透转移将是 $\nu_1 g\cdot B$，需要指出，这时计算后的结果将是相对于阱底势 V_1 的穿透转移值。表 7-4 分别列出了各个势垒高度时及各种能级差时的质子在每秒内的穿透几率和 Boltzmann 因子值。

表 7-4　每秒内质子的穿透机率 $\nu_1 g$

V_0 or ΔE	n	m	V_0 or ΔE	n	m
0.0	13.000	0	1.5	1.978	−24.39
0.1	10.172	−1.626	2.0	0.272	−32.52
0.2	8.975	−3.252	2.5	−1.230	−40.65
0.3	8.070	−4.878	3	−2.588	−48.78
0.4	7.308	−6.504	4	−6.000	−65.04
0.5	6.636	−8.130	5	−7.125	−81.30
1.0	4.000	−16.260	6	−9.045	−97.56

* 计算穿透机率时，$a_0=0.6$Å，$V_0=V_2-V_3$，并以 $\nu_1 g=10^{-n}$ 形式表示结果；Boltzmann因子 $B=\exp(-\Delta E/kT)$，计算时取 $T=310$K，计算结果以 $B=10^m$ 形式表示，能量单位以eV表示。

我们通过上面的计算，可以得到质子转移的数量级概念。现在再回过来讨论 DNA 中质子的自发转移。前面已讲过，质子在一个碱基里所处位置的变动会导致互变异构体的产生，这是复制时产生突变的一个来源。我们来看碱基对间质子隧道效应引起的情况，在通常情况下，碱基间电荷是平衡的，当一个质子向某一方向穿透时，为了保持总电荷中性，极可能引起相反方向的另一个质子的穿透，即二个质子在相反方向同时穿透的净结果将是

:H→:　:　H:
:←H:　→　:H　:

从而导致碱基转变，即产生了一对互变异构碱基对：

A-T→A*-T*，G-C→G*-C*

如果氢键在这时拆开，则会有一种异构复制的形式产生，即会有一个突变，如图7-21：

我们要利用前面计算的结果来估计这种因质子隧道效应引起的自发突变率有多大。氢键里的质子处于双阱势势场，其二个孤对电子供体分属于嘌呤或嘧啶，而其电荷分布当然是不同的（参见第四章），因此氢键的双阱势不对称。作为一级近似，可用下式计算两个不对称双阱势势阱底的差：

$$V_3-V_1=7.199(q_1-q_2)\ \text{(eV)} \tag{7.27}$$

这里 q_1、q_2 分别是氢键里二个电子对供体的电荷密度，其值取自 Pullman[28] 的 HMO 计算值如图 7-22。由此可得氢键里的阱底势差，“+”表示最低的阱在键左边，“−”表示在右边，能量后的数字系 $T=310$ K 时，最小穿透能级时所对应的 Boltzmann 因子：

A-T	G-C：
+1.886eV；10^{-31}	−1.742eV；10^{-28}
−2.606eV；10^{-42}	+0.957eV；$10^{-15.5}$
	+1.864eV；10^{-30}

前面已提到，Boltzmann 因子对势垒高度 V_0 很敏感，因此它的值要正确才能得出有意义的结果，不过我们可在这里作一些合理的估计。事实上 a_0 远小于 1 Å，例 $a_0\approx0.6$Å。从上面列出的 Boltzmann 因子可以看出，除了 G-C 对中间的氢键外，其余氢键里的质子转移可能性极小，即 DNA 中的遗传密码除了一个可能的例外以外其余是十分稳定的。因此着重考虑 G-C 对之中间氢键的质子转移。取 $b_1\approx0.3$Å，根据式 (7-23)，在一个复制循环 τ 秒里，质子的穿透几率将是

$$\nu_1 g\tau=\tau\times10^{13-9V_0^{1/2}} \qquad (7.28)$$

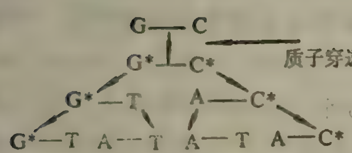


图 7-21 质子穿透引起的突变

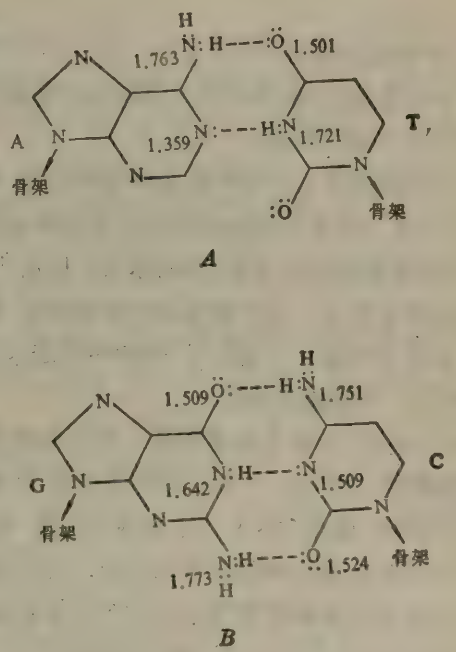


图 7-22 互补碱基对中氢键里电子对供体的 HMO 电荷密度

取 $\tau\approx10^3$sec，即 20 分钟；$V_0=2$ eV，则 $\nu_1 g\tau\approx10^3$，再乘以 Boltzmann 因子 $10^{-15.5}$，则可得在每一代 G-C 对复制中其中间氢键的总势垒高度为 $V_0+(V_3-V_1)=2+0.96=2.96$eV，则其 Boltzmann 因子为 10^{-48}，这实际上表明质子越过势垒的可能性极小。因此，主要因素是穿透。对于 $V_0=1$ eV，每代 G-C 对复制中的质子转移率的数量级将为 $10^{-8.5}$，这时总势垒高度是 1.96eV，其 Boltzmann 因子为 10^{-32}，质子越过势垒可能性也极小。

有意思的是 Pullman[28] 夫妇曾用一种完全不同的处理方法，即通过共振能或离域能来讨论碱基的稳定性，也得出了 G-C 对突变率要高于 A-T 对这一结论。这些结果对于讨论与 DNA 分子碱基成分变化有关的进化理论非常重要。

要说明一下：上述数量级概念导得的过程中，用的是 HMO 计算的电荷，这与 Hückel 参数

选择很有关系。上述结果是初步的。要得到碱基对中氢键计算的满意结果，只有通过解量子力学多体问题以充分考虑质子移动对 σ 和 π 电子极化的影响。才有可能。

再来讨论一下 DNA 中诱导的质子转移。如果由于某些原因使碱基对上的电荷分布变化，例如由于电离辐射、施主-受主反应、质子加到嘌呤碱基N 3 或N 7 孤对电子上，或外电场存在等等，都可以使氢键中的双阱势形状变化，这样质子转移的机率将会增加几个数量级。如果还有些情况使势垒完全消失或降低很多，这样穿透和越过势垒都很容易发生。在这些情况下，质子转移主要向一个方向进行：

$$\begin{array}{ccc} \overrightarrow{\ } & & \\ :\mathrm{H}\cdots: & \rightarrow & :\cdots\mathrm{H}: \\ :\cdots\mathrm{H}: & & :\cdots\mathrm{H}: \end{array}$$

这导致 A-T 型转变成 A^{-}-T^{+}、A^{+}-T^{-}型，或 G-C 型转变成 G^{-}-C^{+}、G^{+}-C^{-}型；而离子型的碱基不能和正常碱基结合，结果可能会引起这对碱基在复制过程中的消失。因此，这种方式的质子诱导转移将导致不可逆突变。

以上介绍的就是 Löwdin 的遗传密码质子-电子对模型。Löwdin的复制平面模型讨论的是“结合错误”，即一种由于环境中的互变异构体对结合到复制的 DNA 中而引起的自发突变，而他的质子-电子对模型讨论的是“异构错误”，即由于 DNA 碱基对中的氢键里的质子转移形成异构体而引起的自发突变。这些模型虽然都较粗糙，但其定性结果却是令人鼓舞的，Löwdin 模型的成功在于量子力学进入了遗传学，使遗传学进入了一个新的层次——量子遗传学。

7.2.1.4 一级微扰模型

Löwdin 的质子-电子对模型是一种不涉及代谢的突变(metabolic independent mutation, MIM)微观模型。Drake[29]等以代谢惰性的 T_4 病毒粒子之 DNA 分别于 0°C和 20°C时放在一种稳定的介质中，使此 DNA 处于细胞外环境，观察到随时间的增加点突变损伤有积累。即存在非诱导的不涉及代谢的转位和转化突变。这一实验结果支持了 Löwdin 的模型。对 Löwdin 模型的批评主要在于该模型所预示的 DNA 复制稳定性太高，另外还有计算表明在 37°C时人类每个细胞每 24 小时内大约有 100 次非诱发的不涉及代谢的突变发生[30,31]。因此 Cooper[32]提出 DNA 氢键内质子迁移的一级微扰模型以从动力学角度来定性地讨论碱基对中的质子迁移；而 Löwdin 模型实质上讨论的是质子转移的“热力学”稳定性。

Cooper 模型的基本思想是 DNA 股间氢键内的质子的近端原子和质子的远端原子因微扰而引起相互作用，微扰提供了相互作用的交换能量，使近端原子变成远端原子，远端原子变成近端原子，即质子相对于这二个原子来说位置是转移了。交换能量减低则质子转移容易发生，因此中心问题是要求处在什么条件下，这一交换能量会降低，即质子转移几率增加。下面我们来介绍 Cooper 的一级微扰模型。

现考虑 DNA 股间质子转移的模型反应：

$$\mathrm{O}\cdots\mathrm{H{-}N}\rightarrow\mathrm{O{-}H}\cdots\mathrm{N}$$

模型氢键的粒子间坐标如图 7-23；该系统的整个势能面如图 7-24，可以看出这种势是分块的模型等能面。

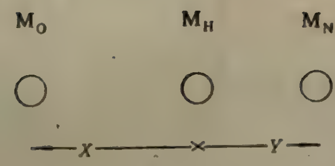


图 7-23 共线氢键系统，x 坐标是从 O 原子延伸至质子 M_H，y 坐标是从 N 原子延伸至质子

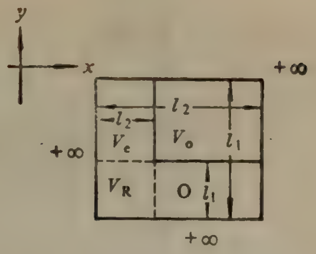


图 7-24 模型氢键体系的位能面

N靠近H的氢键构型(hydrogen bond configuration)(O···H—N)用非相互作用态$|\alpha\rangle$表示，H—N振子的坐标为Y，而氧原子O则沿X坐标振荡。O靠近H的氢键构型O—H···N用非相互作用态$|\beta\rangle$表示，这时远端原子N的振荡则沿y轴方向，而X坐标则是振子O—H的位移。以下用ρ表示态记号。整个位能面由V_e、V_R和V_0三块模型能面组成。V_e是模型反应反应时吸收的能量，V_R是在相互作用区为

$$\begin{Bmatrix}0\leqslant x\leqslant l_2\\0\leqslant y\leqslant l_1\end{Bmatrix}$$

的势能，V_0是双原子振子H—N态$|\alpha\rangle$的结合能。这里l_1、l_2分别是H—N、O—H的方势阱宽度，而L_1、L_2分别是态$|\beta\rangle$、$|\alpha\rangle$里的中心质子至远端原子N及O的距离。即整个位能面可看作是可分的和不可分的成份的叠加：

$$V(XY)=V_\rho(Y)+V_\rho(X)+\Delta V(XY), \tag{7.29}$$

这里

$$V_\alpha(Y)=\begin{cases}\infty, & (Y<0)\\0, & (0\leqslant Y\leqslant l_1)\\V_0, & (l_1\leqslant Y\leqslant L_1)\\\infty, & (Y>L_1),\end{cases} \tag{7.30}$$

$$V_\alpha(X)=\begin{cases}\infty, & (X<0)\\V_e-V_0, & (0\leqslant X\leqslant l_2)\\0, & (l_2\leqslant X\leqslant L_2)\\\infty, & (X>L_2),\end{cases} \tag{7.31}$$

$$V_\beta(Y)=V_\alpha(Y)-V_0, \tag{7.32}$$

$$V_\beta(X)=V_\alpha(X)+V_0, \tag{7.33}$$

$$\Delta V(XY)=\begin{cases}V_R+V_0-V_e, & \begin{Bmatrix}0\leqslant X\leqslant l_2\\0\leqslant Y\leqslant l_1\end{Bmatrix}\\0,\text{其他区域}\end{cases} \tag{7.34}$$

态$|\alpha\rangle$时的二个方势阱如图7-25，态$|\beta\rangle$时的二个方势阱如图7-26，微扰ΔV的势如图7-27表示，从图7-26也可知双原子振子O—H的结合能是V_0-V_e。态$|\rho\rangle$的Hamilton算符H_ρ如下：

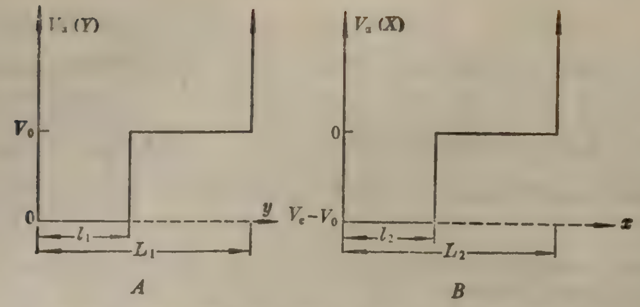


图 7-25 态$|\alpha\rangle$的势，A. 双原子振子 H—N，B. 远端氧原子

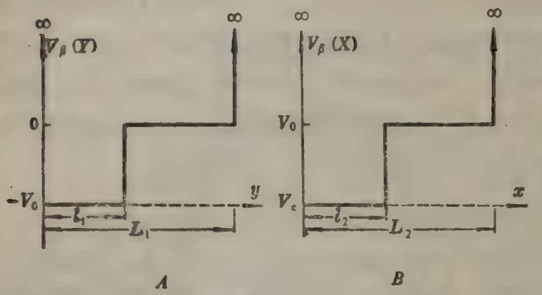


图 7-26 态$|\beta\rangle$的势，A. 远端氮原子，B. 双原子振子 O—H

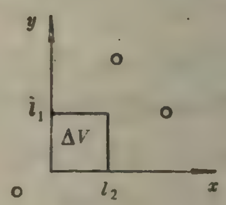


图 7-27 微扰成份的能

$$H_\rho = T_\mu + T_{\mu\rho} + V_\rho(X) + V_\rho(Y) \tag{7.35}$$

这里，$T_{\mu\rho}$是双原子振子(H—N 或 O—H)的动能算子；T_μ是远端原子的动能算子，态$|\alpha\rangle$、态$|\beta\rangle$的双原子振子的约化质量分别如下：

$$\mu_\alpha = \frac{M_H M_N}{M_H + M_N} \tag{7.36}$$

$$\mu_\beta = \frac{M_H M_O}{M_H + M_O} \tag{7.37}$$

态$|\rho\rangle$满足本征值方程：

$$H_\rho|\rho\rangle = E_\rho|\rho\rangle, \tag{7.38}$$

这里 E_ρ是系统的总能量。从式(7.35)至(7.37)可知，若氢键里的端原子质量 M_O和 M_N一样，则$|\alpha\rangle$态及$|\beta\rangle$态将是同一个 Hamilton 量的本征态，将会正交，这一假定对于三粒子氢键的定性模型是允许的。

我们从$|\alpha\rangle$和$|\beta\rangle$间非反应的稳态氢键体系出发来讨论，反应机率$\mathscr{W}_{\alpha\beta}$可表示为：

$$\mathscr{W}_{\alpha\beta} = \frac{2\pi}{\hbar}|\langle\beta|\Delta V|\alpha\rangle|^2\delta(E_\alpha - E_\beta), \tag{7.39}$$

微扰 ΔV 是整个位能面的组成部分，它使体系反应。采用 δ 函数则可使能量保守。在态$|\rho\rangle$时，二个一维粒子体系的联合波函数满足下列方程：

$$\left\{\frac{\hbar^2}{2\mu_\rho}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2}\right) + E_\rho - V_\rho(X) - V_\rho(Y)\right\}\Psi_\rho(XY) = 0, \tag{7.40}$$

这里归一化的质心坐标允许用一个简化质量表示。

束缚态双原子振子的本征函数是

$$\xi_\rho(n|S)=\begin{cases}C_{n\rho}\sin(K_{n\rho}S), & (0\leqslant S\leqslant l)\\ A_{n\rho}\exp(-\gamma_{n\rho}S)+B_{n\rho}\exp(\gamma_{n\rho}S), & (l\leqslant S\leqslant L),\end{cases} \tag{7.41}$$

$S=X,l=l_2$ 及 $L=L_2$ 当 $\rho=\beta$ 时；$S=Y,l=l_1$ 及 $L=L_1$ 当 $\rho=\alpha$ 时；

$$K_{n\alpha}^2=\frac{2\mu_\alpha}{\hbar^2}E_{n\alpha}, \tag{7.42}$$

$$\gamma_{n\alpha}^2=\frac{2\mu_\alpha}{\hbar^2}V_0-K_{n\alpha}^2, \tag{7.43}$$

$$K_{n\beta}^2=\frac{2\mu_\beta}{\hbar^2}E_{n\beta}, \tag{7.44}$$

$$\gamma_{n\beta}^2=\frac{2\mu_\beta}{\hbar^2}(V_0-V_e)-K_{n\beta}^2, \tag{7.45}$$

$$B_{n\rho}=\frac{2\gamma_n\xi_n\exp(-2\gamma_n l)}{4\gamma_n[K_n l-\sin(2K_n l)]\xi_n\{1+[4L-\exp(2\gamma_n L)]\exp(2\gamma_n L)\}} \tag{7.46}$$

$$A_{n\rho}=-B_{n\rho}\exp(2\gamma_{n\rho}L), \tag{7.47}$$

$$C_{n\rho}=2B_{n\rho}\left[\frac{\gamma_{n\rho}}{K_{n\rho}\cos(K_{n\rho}l)+\gamma_{n\rho}\sin(K_{n\rho}l)}\right]\times\exp(\gamma_{n\rho}l)。 \tag{7.48}$$

式(7.46)里态 ρ 的下标压缩掉了，ξ_n 表示 $K_{n\rho}[K_{n\rho}\cos(K_{n\rho}l)+\gamma_{n\rho}\sin(K_{n\rho}l)]^2$，$\mathscr{L}$ 表示差 $L-1$。以式(7.41)代入下式可得

$$\int_0^L\xi\xi^*\mathrm{d}S=1 \tag{7.49}$$

$B_{n\rho}$。$E_{n\alpha}$、$E_{n\beta}$ 分别是双原子分子 H—N 和 O—H 的振动能量。束缚态能量值可由量子条件：

$$\tan(K_{n\rho}l)=-\left[\frac{K_{n\rho}}{\gamma_{n\rho}}\right]\left\{\frac{\exp(2\gamma_{n\rho}[L-l])-1}{\exp(2\gamma_{n\rho}[L-l])+1}\right\}, \tag{7.50}$$

求得，这一条件是由在 $S=l$ 时对 $\xi_\rho(n|S)$ 的连续性要求而来。远端的振荡的原子的波函数为：

$$\chi_\rho(n|S)=\begin{cases}c_{n\rho}\sin(\overline{\gamma}_{n\rho}S), & (0\leqslant S\leqslant l)\\ b_{n\rho}\sin(\overline{K}_{n\rho}S+\eta_{n\rho}), & (l\leqslant S\leqslant L),\end{cases} \tag{7.51}$$

这里，$S=X,l=l_2$ 及 $L=L_2$ 当 $\rho=\alpha$ 时；$S=Y,l=l_1$ 及 $L=L_1$ 当 $\rho=\beta$ 时。其波数、振幅、相位 $\eta_{n\rho}$ 分别计算如下：

$$\overline{K}_{n\alpha}^2=\frac{2\mu_\alpha}{\hbar^2}E_\alpha-K_{n\alpha}^2, \tag{7.52}$$

$$\overline{\gamma}_{n\alpha}^2=\overline{K}_{n\alpha}^2-\frac{2\mu_\alpha}{\hbar^2}(V_e-V_0), \tag{7.53}$$

$$\overline{K}_{n\beta}^2=\frac{2\mu_\beta}{\hbar^2}(E_\beta-V_e)-K_{n\beta}^2, \tag{7.54}$$

$$\overline{\gamma}_{n\beta}^2=\overline{K}_{n\beta}^2+\frac{2\mu_\beta}{\hbar^2}V_0, \tag{7.55}$$

$$\eta_{n\rho}=-\overline{K}_{n\rho L}, \tag{7.56}$$

$$b_{n\rho}=\frac{4\overline{K}_n\bar{\xi}_n}{2\Xi\{K_nl-\sin(2K_nl)\}\bar{\xi}_n\{2\overline{K}_nL-\gamma_n\}} \tag{7.57}$$

这里，式(7.57)的下标 ρ 被压缩：$\bar{\xi}_n$ 代表$[\sin(\bar{\gamma}_{n\rho}l)+\bar{\gamma}_{n\rho}\cos(\bar{\gamma}_{n\rho}l)]^2$；$\Xi$代表$[\sin(\overline{K}_{n\rho}l+\eta_{n\rho})+\overline{K}_{n\rho}\cos(\overline{K}_{n\rho}l+\eta_{n\rho})]^2$，$\gamma_n$ 代表 $\sin[2(\overline{K}_{n\rho}L+\eta_{n\rho})]-\sin[2(K_{n\rho}l+\eta_{n\rho})]$，$\mathscr{L}$ 代表 $L-1$。式(7.51)中的 $c_{n\rho}$ 则如下式计算：

$$c_{n\rho}=b_{n\rho}\left\{\frac{\sin(\overline{K}_{n\rho}l+\eta_{n\rho})+\overline{K}_{n\rho}\cos(\overline{K}_{n\rho}l+\eta_{n\rho})}{\sin(\bar{\gamma}_{n\rho}l)+\bar{\gamma}_{n\rho}\cos(\bar{\gamma}_{n\rho}l)}\right\}, \tag{7.58}$$

式(7.57) 中的 $b_{n\rho}$ 可用归一化条件$\int_0^L\chi\chi^*\mathrm{d}S=1$ 来求得。因此在相互作用区域内：

$$\begin{Bmatrix}0\leqslant X\leqslant l_2\\0\leqslant Y\leqslant l_1\end{Bmatrix}$$

态$|\rho\rangle$的波函数则如下式：

$$\Psi_\rho(n|ss')=c_{n\rho}C_{n\rho}\sin(\bar{\gamma}_{n\rho}S)\sin(K_{n\rho}S'), \tag{7.59}$$

这里，当$|\alpha\rangle$态时，S、S'分别为X和Y；当$|\beta\rangle$态时 S、S' 分别为Y和X。

我们的目的是要求出转移率增加时的条件。从式(7.39)知，要求出转移率 $\mathscr{W}_{\alpha\beta}$，必须知道在相互作用区域内态$|\rho\rangle$的波函数。这就是上面冗长的推导的目的。现在，我们可以将波函数代入式(7.39)，来解与转移率有关的条件。

在把波函数式(7.59)代入式(7.39)后，可得：

$$\mathscr{W}_{\alpha\beta}=\frac{2\pi}{\hbar}\left|c_{n\alpha}c_{n\beta}C_{n\alpha}C_{n\beta}\Delta VF(l_1)F(l_2)\right|^2, \tag{7.60}$$

这里，

$$F(l_1)=\int_0^{l_1}dY\sin(K_{n\alpha}Y)\sin(\bar{\gamma}_{n\beta}Y) \tag{7.61}$$

$$F(l_2)=\int_0^{l_2}dX\sin(K_{n\beta}X)\sin(\bar{\gamma}_{n\alpha}X). \tag{7.62}$$

积分后，式(7.60)则为：

$$\begin{aligned}\mathscr{W}_{\alpha\beta}=&\frac{2\pi}{\hbar}\left|c_{n\alpha}C_{n\alpha}c_{n\beta}C_{n\beta}\Delta V\right|^2\\&\times\left\{\left[\frac{\sin[(K_{n\beta}-\bar{\gamma}_{n\alpha})l_2]}{2(K_{n\beta}-\bar{\gamma}_{n\alpha})}-\frac{\sin[(K_{n\beta}+\bar{\gamma}_{n\alpha})l_2]}{2(K_{n\beta}+\bar{\gamma}_{n\alpha})}\right]\right.\\&\left.\times\left[\frac{\sin[(K_{n\alpha}-\bar{\gamma}_{n\beta})l_1]}{2(K_{n\alpha}-\bar{\gamma}_{n\beta})}-\frac{\sin[(K_{n\alpha}+\bar{\gamma}_{n\beta})l_1]}{2(K_{n\alpha}+\bar{\gamma}_{n\beta})}\right]\right\}^2,\end{aligned} \tag{7.63}$$

我们现就可以根据式(7.63)来分析氢键从态$|\beta\rangle|\alpha\rangle$转变为态$|\alpha\rangle|\beta\rangle$的跃迁率。利用等式：

$$\sin(a\pm b)=\sin a\cos b\pm\sin b\cos a, \tag{7.64}$$

可以把式(7.63)中大括号里对应于 $F(l_2)$的项化为：

$$F(l_2)=\frac{\{[K_{n\beta}\cos(K_{n\beta}l_2)]^2+[\bar{\gamma}_{n\alpha}\sin(K_{n\beta}l_2)]^2\}^{1/2}\sin(\bar{\gamma}_{n\alpha}l_2-h_2)}{\bar{\gamma}_{n\alpha}^2-K_{n\beta}^2} \tag{7.65}$$

这里，

$$h_2 = \tan^{-1}\left\{\frac{\overline{\gamma}_{n\alpha}\sin(K_{n\beta}l_2)}{K_{n\beta}\cos(K_{n\beta}l_2)}\right\} \tag{7.66}$$

现在来看式(7.65)的分母：

$$\overline{\gamma}^2_{n\alpha} - K^2_{n\beta},$$

这里 $\overline{\gamma}_{n\alpha}$ 是态 $|\alpha\rangle$ 的远端氧的波数，$K_{n\beta}$ 是态 $|\beta\rangle$ 的O—H 振子的 波数。 可以 看出当波 数差变小，即 $|\alpha\rangle$ 态中远端氧沿 X 轴的振荡波数和 $|\beta\rangle$ 态中 O—H 振子里 的氧沿 X 轴 的振 动波 数接近时，则 $\mathscr{W}_{\alpha\beta}$ 变大，即这时氢键从 $|\beta\rangle$ 态变成 $|\alpha\rangle$ 态的几率增加。同样也可以把式(7.63) 中 大括号里对应于 $F(l_1)$ 的项经过变换得到分母为

$$\overline{\gamma}^2_{n\beta} - K^2_{n\alpha}$$

的项。这里 $\overline{\gamma}_{n\beta}$ 是 $|\beta\rangle$ 态中远端氮的波数，$K_{n\alpha}$ 是振子 H—N 的波数。同样也能看出，这一项中分母变小，即波数差趋近于零时对 $\mathscr{W}_{\alpha\beta}$ 的贡献，这时氢键从 $|\alpha\rangle$ 态变成 $|\beta\rangle$ 态的几率增加。

从式(7.42)、(7.44)、(7.52)、(7.53)、(7.54)、(7.55)、还可知差

$$\overline{\gamma}^2_{n\beta} - K^2_{n\beta}$$

包括振子的振动能量和远端原子的能级之差。

所以我们得出质子转移率增加的条件是：氢键中端原子在两个态里的波数要“相似”；亦即振子振动能量要大而远端原子的能级要低。而微扰 $\Delta V(XY)$ 是描述氢键里两个端原子的 排 斥相互作用：微扰加于体系时，两个端原子相互作用，近端原子和远端原子发生能量交换：和质子束缚在一起的近端原子因排斥作用获得能量而被激发至远端，远端原子失去能量因 吸引 作用而被束缚于质子近端。这一交换所需的能量减小时，模型反应的反应率即质子转移率增加，而这一能量是与振子的振动能量及远端原子的能级之差相关的。在总能量固定时，振子的 振动能 量增加时，则质子转移率增大。这些就是 Cooper 模型的中心结论。

问题是明显的，在体内，DNA生物功能的发挥是不能用固定的、具刚性晶体结构的双螺旋模型来解释的。因此，在体内的 DNA 股间氢键里二个端原子的距离会随生理条件的不同而变化。在 0～37°C范围里，温度升高使热振动加强，这会使质子转移的反应率增加 10 倍以上[33]。这 些情况，反映了体内 DNA 的“柔性”，可用 Cooper 模型说明。在 Cooper 模型里把氢键原 子 间 的运动看作双原子振子和远端原子运动的叠加，双原子振子次能量上可及的本征态振动；而远端原子的运动又可看作是其在氢键里的原子振动和宏观振荡运动的叠加，这一宏观振荡即 是把 两股 DNA 视作经典的“类粒子”间的相互作用。因此，远端原子和氢键里质子间的距离随时间而涨落。当 DNA 某部位的振子的振动运动和远端原子的振荡运动“波数相似”时或者该部位的振子振动能加强而远端原子能级下降时，将会使质子发生转移的 $\mathscr{W}_{\alpha\beta}$ 增加，这个部位即可能对应于 所谓的对非诱导遗传变化极为敏感的“热点”突变部位。如果 DNA 螺旋某部位由于某些条件使 $\mathscr{W}_{\alpha\beta}$ 下降，则这部位的正常的遗传质子-电子对码就可稳定。这些说明了为 什 么 DNA 中某些点易突变。Cooper 的微扰模型从定性上讨论了内能对质子转移率的影响，这一转移率即正比于模型反应的反应几率。

7.2.2 基因表达及其控制的数理模型

7.2.2.1 DNA 转录的物理模型

转录是基因表达的第一步。DNA 转录过程中，RNA 聚合酶起着关键的作用。用物理方法来研究转录机制的主要困难是 RNA 聚合酶极少“停”在 DNA 上[34]，因此在 RNA 合成时，模板构象的转变仅仅在 RNA 合成的活性部位起作用，而整个 DNA 的长分子不受影响。所以研究的注意力往往就放在 RNA 聚合酶和那极短的一段 DNA 所形成的复合物上。Rüger[34] 认为与 RNA 聚合酶结合的DNA 部分不到三个螺距(大约 30 个核苷酸对)，并发现转录完后这部份仍保留完整。曾合成过一些上述的复合物.这些实验的模型复合物有着聚合酶的基本功能，可以合成 15～20 个核苷酸的一段 RNA 分子。Florentiev and Ivanov[35]曾提出 RNA 聚合酶的作用模型，如图 7-28：

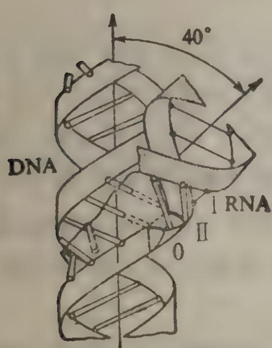


图 7-28 RNA 聚合酶作用模型，图中表示了 DNA 模板及生长着的 RNA 端部

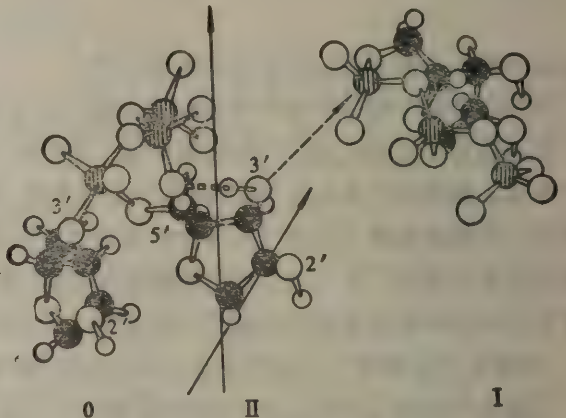


图 7-29 核苷酸结合的二步示意图；左边是 RNA 生长部位处的二个已结合的核糖核苷酸，右边是新形成的要结合上去的核苷三磷酸酯；0 是核苷酸，其与 II 形成一共价键，此时还未释放出焦磷酸酯

按照 Florentiev 和 Ivanov 模型.当 RNA 链增长时，核糖为 3′-羟基和攻击核苷酸之 5′-磷酸酯上的基团结合，而 RNA 碱基的顺序则由 DNA 碱基的顺序决定。已经知道，磷酸酯键的断裂过程有着一个中间态.这时 P 原子的五个共价键处于一个三角双锥的构型。因此，这个模型里核苷酸的加成反应采用二步表示，如图 7-29。中间态(II)的形成需要能量，而焦磷酸酯的分离则可释放能量。这模型考虑了二步之间的过渡。该模型理论化的地方是：DNA 模板以 A 型式(不是通常的 B 型式)参与 DNA 的解旋和 DNA-RNA 杂化双螺旋的形成。这意味着 A 型式在模板的反应体系活性部位邻近区域是稳定的。这一假定和一系列事实相符合。在聚合酶起作用时，B 型—→A 型的局部转化可以用使水分子从酶结合的 DNA 区域脱去从而有利于核苷酸的缩聚的方法来实现。在这个模型里，当互补碱基对链的氢键断裂之后，一个 DNA 链上的碱基从 A 型式中糖苷的沟道里移去时不存在空间障碍，而第二条链的构象并不起变化。在二个阶段过渡时，二个核苷酸同时参与，从图 7-29 可见，一个核苷酸处于状态I，另一个处于状态II；这二个核苷酸都

和 DNA 核苷酸配对。在模型里，双螺旋杂化体的轴和模板 DNA 双螺旋的轴间夹角为 40°。这样，每一步新碱基从 DNA 模板上"抽出"，而脱氧核糖核苷酸在和其互补的核糖核苷酸，失去焦磷酸酯后，回到原来的 DNA 上并重新恢复氢键。杂化体双螺旋的轴象绕螺旋一样沿着 DNA 模板的轴向前移，酶也同时换到新的地方。由于已知在无底物时 RNA 聚合酶只是很弱地和 DNA 结合。另一方面、又发现在 RNA 合成时，有一个非分离的 DNA-酶复合物形成[36]。因此，根据 Florentiev 和 Ivanov 模型，这一复合物是以两个相互作用的核糖核苷三磷酸酯的焦磷酸酯上的基团和聚合酶亚单位上的阳离子基团间的键而稳定的。这些三磷酸酯的碱基和 DNA 的碱基形成 Watson-Crick 对。核苷酸 I 和 II 跟核苷酸 0 形成一个中间杂化体区域；在任何时刻，有二个核糖核苷三磷酸酯起作用，因此这模型解释了强的酶结合。模型表示了核苷酸 0 和 II 的糖-磷酸酯部分的构象，并对核苷酸 I 的构象也作了假定；当酶沿 DNA 一个个碱基替换时，则构象变化到有利于反应。这一模型还考虑了聚合酶的亚单位结构和每一步里结构的重排；因此从本质上来说，这模型考虑了电子-构象相互作用。

Riley[37]提出了另一种模型。Riley 的模型假定转录时 DNA 双螺旋不解旋；RNA 在 DNA 的较宽广的沟道内增长，这是因为在核糖核苷酸和 DNA 碱基对间有一种专一的立体化学相互作用。在中间态阶段，由二个 DNA 碱基和一个 RNA 碱基形成一三联体；这模型不考虑酶的作用。

在 Florentiev 和 Ivanov 模型及 Riley 模型提出之后不久，实验数据肯定了解旋模型；并且圆二色性方法(CD)分析表明，在 mRNA 合成时，有 A 型的 DNA 或 A 型的 DNA-RNA 杂化体出现[40]。因此 Florentiev 和 Ivanov 模型较接近实际。

7.2.2.2 密码子

1) 密码子的信息论证明　　我们在前面已讲过，mRNA 模板上的碱基顺序将决定所合成的蛋白质分子的氨基酸顺序。现在要问为什么密码子是三联体，即由三个碱基组成。

已经知道生命蛋白质分子约有 20 种氨基酸组成。我们也已证明为什么 DNA 的碱基要有4种。现在要来证明：为什么密码子由三个碱基组成[16]。

根据信息论，每个氨基酸所载有的信息是：

$$i_a = -\log_2\frac{1}{20} = 4.322\ \text{bit} \tag{7.67}$$

而每个碱基所载有的信息是：

$$i_b = -\log_2\frac{1}{4} = 2.000\ \text{bit} \tag{7.68}$$

因此每个氨基酸所需要的最少的碱基数目 n 是：

$$i_a = n \cdot i_b \tag{7.69}$$

$$n = \frac{i_a}{i_b} = 2.161 \tag{7.70}$$

自然，碱基数目不能是非整数，因此将 n 调整为 3，即与一个氨基酸对应的密码子至少要由 3 个碱基组成，即三联体。

事实上，mRNA 里共有 4 种碱基，每次取 3 个作不同排列，共有 $4^3=64$ 种排列，对于 20 种氨基酸来讲已是足足有余，因此密码子必然会有简并情况存在。

2）广谱遗传密码(generalized genetic code)　已经发现：在离体条件下，密码子有多义性，即一个密码子可对应于 n 个氨基酸；在某些细菌体系里，密码子也有多义性[38,39]。另外，早已知道在蛋白质合成起始时，具有多义现象，即密码子 AUG、GUG 可作为起始密码子，而又分别作为 Met 和 Val 的密码子。对此，不少生物学家认为遗传密码事实上是种“映照”(mapping，即单值的对应)，而离体或体内研究中产生的多义性只是表示翻译系统并不是不会发生错误的。但 Gatlin[45]认为遗传密码的多义性存在表示遗传密码是一种“关系”(relation，即多值对应)而不是“映照”。他的假说是：这种关系所内含的多义性可以为翻译密码子时的生物环境条件所移去。Gatlin 用上述假说来解释蛋白质合成起始的多义现象，他认为在这个例子中“起始”就是这种生物环境条件，在“起始”时，AUG的多义性即被除去。Findley[46]等采纳了 Gatlin 的观点，认为多义性是遗传密码整体的一部分，是正常的；并根据群论理论提出了广谱遗传密码理论。

下面就来介绍 Findley 的广谱遗传密码*

先看标准密码(standard genefic code，SGC)。遗传密码是种语言，它是基于 4 个核酸碱基集 $\mathscr{B}=\{U,A,C,G\}$ 的语言。集 $\mathscr{B}$ 的元素是嘧啶(U、C)和嘌呤(A、G)碱基。模板 mRNA 上碱基的线性顺序决定了其对应的蛋白质分子中的氨基酸顺序。三个碱基联合体组成一对应的蛋白质的密码子。

密码子集 $\mathscr{C}$ 由集 $\mathscr{B}$ 中的所有可能的有序列的三元素联合体组成，即 $\mathscr{B}$ 和 $\mathscr{C}$ 以下列积相联系：

$$\mathscr{C}=\mathscr{B}\times\mathscr{B}\times\mathscr{B}. \tag{7.71}$$

因此，$\mathscr{C}$ 含有 $4^3=64$ 个密码子。还要考虑一个集 $\mathscr{A}$，由 20 个氨基酸和操作子 TC(终止密码子)组成。根据目前的生物学理论，$\mathscr{C}$ 和 $\mathscr{A}$ 以映照相关联：

$$f:\mathscr{C}\to\mathscr{A}. \tag{7.72}$$

这一映照对应是以化学的及遗传学的方法决定的，如下表 7-5，以下将称式(7.72)为标准遗传密码(SGC)。

表 7-5 即规定了 f 的操作：将 $\mathscr{C}$ 映照至 $\mathscr{A}$ 上。映照的投影性质定义了 SGC 的简并性。在某些条件下，SGC 出现多义性。大多数的多义性是在离体条件下产生的，这时体系所处的实验环境条件不同于体内环境。表 7-6 是这些多义性密码子的例子。

根据 Gatlin 的观点，遗传密码是一种“关系”而不是“映照”。因此要重新考虑集 $\mathscr{C}$ 和 $\mathscr{A}$ 的对应相关。这就是下面要讨论的广谱遗传密码。

把多义性看作是遗传密码的一个正常反映，则需将遗传密码视作是一关系 $\mathscr{R}$，此处：

$$\mathscr{R}\subset\mathscr{C}\times\mathscr{A} \tag{7.73}$$

我们无法预先知道 $\mathscr{R}$ 是否为 $\mathscr{C}\times\mathscr{A}$ 的恰当的或不恰当的子集，这需通过一些推理去说明。在

* 这里要说明，第一节基因复制中讨论的遗传密码是指 DNA 碱基互补对；这里讨论的是 mRNA 碱基顺序，即密码子。

表 7-5　标准遗传密码

UUU, UUC } Phe UUA, UUG } Leu	UCU, UCC, UCA, UCG } Ser	UAU, UAC } Tyr UAA, UAG } TC	UGU, UGC } Cys UGA TC UGG Trp
CUU, CUC, CUA, CUG } Leu	CCU, CCC, CCA, CCG } Pro	CAU, CAC } His CAA, CAG } Gln	CGU, CGC, CGA, CGG } Arg
AUU, AUC, AUA } Ile AUG Met	ACU, ACC, ACA, ACG } Thr	AAU, AAC } Asn AAA, AAG } Lys	AGU, AGC } Ser AGA, AGG } Arg
GUU, GUC, GUA, GUG } Val	GCU, GCC, GCA, GCG } Ala	GAU, GAC } Asp GAA, GAG } Glu	GGU, GGC, GGA, GGG } Gly

表 7-6　离体条件下的密码子多义性

氨基酸	密码子
Asp	UAG, CAG, GAG, UGA, GGA, CUA
Tyr	AGG, GGG, GUG, UGU, UGG, CAU
Arg	CCG, GCG, GGC, UGA
Cys	AGU, CGU, UGA, AGC
Lys	CUA, UAA, AAU
Gly	UGG, CGG, AGG
Val	UGU, CGU
Thr	CAC, AAC
His	UCA, CCA
Gln	CUA, UGA
Glu	UGA, GGA
Leu	UUU
Ile	UUU
Ala	AGC
Phe	AGG
Met	AUA
Asn	UGA

生物环境条件下，可选得一子集使其表现映照特征。然而，多义性说明了它是随生物环境条件而变化的，因此有几个映照：

$$(\mathscr{R}, C_i) = f_i;\ i \in I, \tag{7.74}$$

这里

$$f_i : \mathscr{C} \to \mathscr{A}. \tag{7.75}$$

式(7.74)中，C_i 表示某一特定的生物环境条件(例如，C_1 表示“正常”蛋白质合成时的环境，则 $(\mathscr{R}, C_1) = f_1 \equiv f$)；$I$ 是某一尚未定的但一定是有限的指标集；记号$(\mathscr{R}, C_i)$表示从 $\mathscr{R}$ 中选择一

个映照 f_i。因此广谱遗传密码 GGC 则是所有映照的一个类，至此我们得到了广谱遗传密码的一套形式：式(7.74)至式(7.76)：

$$\{f_i, i\in I \mid f_i: \mathscr{C}\to\mathscr{A}\} \tag{7.76}$$

由于要去研究决定遗传密码映照的生物环境条件的本质是很困难的，所以Findley 的方法是列举支持GGC 假说的事实来进行推理。从式(7.74)出发来考察决定 $\mathscr{R}$ 的约束条件，并认为如果某种约束条件确实存在，则 $\mathscr{R}$ 将是 $\mathscr{C}\times\mathscr{A}$ 的一个恰当子集，并且此约束也将适用于每一个映照 f_i。

先来看 SGC 的约束条件。众所周知，SGC 的简并性存在一种对称性约束。观察表 7-5 可以发现对应于同一氨基酸的几个密码子之间有一简单关系：即密码子第三位上碱基的变化，特别是U变C、A变G，不影响其所对应翻译的氨基酸。我们可以把这种关系表达成下列式子，即将 $\mathscr{C}$ 作一特殊的分解，成为 4 个成对非联的子集，每一子集的基数是 16：

$$\mathscr{C}_k \equiv \{c_{ijk}\in\mathscr{C} \mid i,j\in\mathscr{B}\};\ k\in\mathscr{B}, \tag{7.77}$$

这里 c_{ijk} 上的指标 i,j,k 表示密码子上的第一、第二、第三个碱基，式(7.77)将 $\mathscr{C}$ 分配至 4 个子集，每一子集只会有第三个碱基相同的密码子，并可被 f 映照成 $\mathscr{A}$，而其映像则可用 $f(\mathscr{C}_k)$ 表示。这种 $\mathscr{C}$ 的分解(以下用 D_0 表示这种分解)，及其映像如表 7-7：

表 7-7　按照 $\mathscr{C}_k\equiv\{c_{ijk}\in\mathscr{C} \mid i,j\in\mathscr{C}\}$ 的 $\mathscr{B}$ 分解

C_U	$f(C_U)$	C_C	$f(C_C)$	C_A	$f(C_A)$	C_G	$f(C_G)$
UCU, AGU	Ser	UCC, AGC	Ser	UCA	Ser	UCG	Ser
CGU	Arg	CGC	Arg	AGA, CGA	Arg	AGG, CGG	Arg
CUU	Leu	CUC	Leu	CUA, UUA	Leu	CUG, UUG	Leu
GCU	Ala	GCC	Ala	GCA	Ala	GCG	Ala
GUU	Val	GUC	Val	GUA	Val	GUG	Val
CCU	Pro	CCC	Pro	CCA	Pro	CCG	Pro
GGU	Gly	GGC	Gly	GGA	Gly	GGG	Gly
ACU	Thr	ACC	Thr	ACA	Thr	ACG	Thr
UUU	Phe	UUC	Phe	CAA	Gln	CAG	Gln
UAU	Tyr	UAC	Tyr	AAA	Lys	AAG	Lys
UGU	Cys	UGC	Cys	GAA	Glu	GAG	Glu
CAU	His	CAC	His	UAA, UGA	TC	UAG	TC
GAU	Asp	GAC	Asp			UGG	Trp
AAU	Asn	AAC	Asn				
AUU	Ile	AUC	Ile	AUA	Ile	AUG	Met

考察表 7-7，立即可发现

$$f(\mathscr{C}_U)=f(\mathscr{C}_C);\qquad f(\mathscr{C}_A)\neq f(\mathscr{C}_G). \tag{7.78}$$

式(7.78)里的不等式完全是因奇数简并级的密码子造成的，对于偶数简并级的密码子则存在某种简单关系。把具有 i 重简并度的映照至某一 $\mathscr{A}$ 元素上的密码子集以下列记号表示：

$$\{d_1^{(a,i)}, d_2^{(a,i)}, \cdots, d_i^{(a,i)}\}, \tag{7.79}$$

这里角标(a,i)分别表示所讨论的 $\mathscr{A}$ 的元素及其简并度；而足标用来区别各个密码子，明显的有：

$$f(d_1^{(a,i)}) = f(d_2^{(a,i)}) = \cdots = f(d_i^{(a,i)}) = a. \tag{7.80}$$

首先考虑 2 级简并度。从表 7-7 可以得到结论：

或 $$f(d_1^{(a,2)}) \in f(\mathscr{C}_U) \Longleftrightarrow f(d_2^{(a,2)}) \in f(\mathscr{C}_C). \tag{7.81 a}$$

或 $$f(d_1^{(a,2)}) \in f(\mathscr{C}_A) \Longleftrightarrow f(d_2^{(a,2)}) \in f(\mathscr{C}_G). \tag{7.81 b}$$

式(7.81 a、b)定义了二重简并密码子对成员之间的一一对应。如果式(7.81 a) 和 (7.81 b)同时成立，则可定义 4 级简并度；因此这要求在表 7-7 中，每一列中必须有一个并且只有一个 4 级简并度集中的密码子。还可以推广至 6 级简并度；这时对六个密码子中的四个来说，式(7.81 a)和式(7.81 b)同时成立，而对其余二个，则或者式(7.81 a)成立，或者式(7.81 b)成立。

由式(7.81 a、b)定义的一一对应的性质是从分解 D_0 演绎出来的。对于任何密码子对，只要其满足式(7.81 a、b)之其中一式，则该密码子对的前二个碱基相同，而第三个碱基则不同，但一定是嘧啶或嘌呤。这就是 SGC 内的对称性，这里把它叫作偶简并度约束(even-order degeneracy constraint, EODC)；但对奇简并度的密码子，这种对称性不完全成立。下面来讨论这一情况。

考虑二套三度简并密码子，一套映照后得 Ile，另一套映照后得 TC。有二个对应于 Ile 的密码子满足 EODC 和式(7.81 a)，因此这二个密码子可规定为二度简并；剩下的第三个是 AUA，其映像是 f(AUA)＝Ile*。TC 密码子较特别，UAA、UAG 对和 UGA、UAG 对都满足式 (7.81 b)，而第一对还满足 EODC，因此 UAA、UAG 对可规定为二度简并密码子对；剩下的第三个密码子是 UGA，其映像是 f(UGA)＝TC*。最后再来看二个非简并的密码子，一个映照后得 Met，另一个映照后得 Trp。其实对应于 Ile*和 TC* 的密码子也可看作是非简并的。结果发现这 4 个密码子各有一对满足 EODC：(TC*, Trp)、(Ile*, Met)*。因此，可以作一推断：这二对密码子本身是二度简并的。事实上，离体密码子的翻译结果(见表 7-6)，在一定条件下，密码子对(Ile*, Met)是二度简并的，并且它们映照后得 Met。反过来讲，如果认为映照 f 是唯一确切的话，当然这些密码子对就要被视作异常，并看作是在遗传密码的进化过程中产生的关于对称性的小偏差。

根据上面对 SGC 的讨论，可以总结出如下一些规则，考虑下列二个集：

$$F_1 = \{(d_1^{(a,2)}, d_2^{(a,2)}) \in \mathscr{C} \times \mathscr{C} \mid f(d_1^{(a,2)}) \in f(\mathscr{C}_U) f(d_2^{(a,2)}) \in f(\mathscr{C}_C)\}, \tag{7.82a}$$

$$F_2 = \{(d_1^{(a,2)}, d_2^{(a,2)}) \in \mathscr{C} \times \mathscr{C} \mid f(d_1^{(a,2)}) \in f(\mathscr{C}_A) f(d_2^{(a,2)}) \in f(\mathscr{C}_G)\}. \tag{7.82b}$$

对于映照 f (不包括奇数级简并度)，密码子对二度简并的充分和必要条件是：它应是 F_1或是F_2的一个元素。对于四度简并密码子集，四个密码子四简并的充分和必要条件是：一密码子对应是 F_1 的一个元素且另一密码子对应是 F_2 的一个元素。同样，对六度简并密码子集，六个密码子六度简并的充分和必要条件是：二个组成一密码子对，其应是 F_1的一个元素且另外二个组成一

---

* 这里用映像表示其相应的密码子。

密码子对，其应是 F_2 的一个元素；然后最后二个密码子组成一密码子对，其必须是 F_1 或是 F_2 的一个元素。

对于奇数简并度密码子、应用 EODC 时关于严格的对称性，有一些微小的偏差。我们得到了这些规律之后，就可以来讨论广谱遗传密码的对称性。

先要寻找$\mathscr{C}$的另外一些分解，而这些分解具有类似于D_0的在映照 f 下的特征，也就是说要寻找$\mathscr{C}$的分解存在的可能性，这些非 D_0的分解可以构成集 F_1、F_2。

我们的出发点还是式(7.71)，但此时用群论的乘法来处理。对$\mathscr{B}$要定义一个群结构。$\mathscr{C}$ 变成群 $\mathscr{B}$ 的元素三重积的集合，并且在所定义的乘法规则下，当去掉重复元素而又再形成一个集时，$\mathscr{C}$就变得与$\mathscr{B}$等价，即当$\mathscr{B}$是一个群时，映照到群$\mathscr{B}$的元素之一上的$\mathscr{C}$元素的数

$$\mathscr{B}\times\mathscr{B}\times\mathscr{B}=\mathscr{B}. \tag{7.83}$$

且可由$\mathscr{C}$的基数被除于群$\mathscr{B}$的阶而得。因素$\mathscr{B}$的基数是 4，故任何由$\mathscr{B}$形成的群的阶是 4。从而映照到任何一个群元素上的有序的三联体的数是 $64/4=16$。将这些成对的非联的 $\mathscr{C}$ 的子集标记为 $\mathscr{C}_k, k\in\mathscr{B}$，足标表示在群乘法定义下子集密码子映照其上的 $\mathscr{B}$ 的元素。这些记号与前面讨论 SGC 时的类似，但要注意，这里定义的子集 $\mathscr{C}_k$ 的元素的总数不同于 SGC 的情况。

现因$\mathscr{B}$必须是一 4 阶群，故其只有两种可能性，均是 Abel 的：Klein 4-群 V 和 4 阶循环群 Z_4。V 的表示是

$$V:\langle a,b;a^2,b^2,ab=ba\rangle. \tag{7.84}$$

因此群元素是 $V\equiv\{1,a,b,c(\equiv ab)\}$，此处 1 代表恒等元素。$Z_4$ 的表示是

$$Z_4:\langle d;d^4\rangle \tag{7.85}$$

则 $Z_4\equiv\{1,d,e(\equiv d^2),f(\equiv d^3)\}$。

为求 $\mathscr{C}$ 的分解，要根据$\mathscr{B}$对 V 或 Z_4 按次序构作一同素异型式*。一般对于每一抽象群，有 24 个有序的同素异型式，故总的有 48 个。但对 V 只有 4 个不同的有序同素异型式，对 Z_4只有12 个。而当这些有序同素异型式用于导出的分解时，只有 7 个不同的分解 D_1—D_7，如表 7-8。

有了这些分解之后，就可以对它们的对称性质进行讨论；结果发现，只有三种分解 D_1—D_3 在关于f下的对称特征方面完全类同于原来的分解D_0，即这些分解可以导得和D_0一样的隐合条件。但还有四种分解D_4—D_7在关于f下的对称性方面不完全，即它们的4阶、6阶简并度的EODC是模糊的。我们在下面来较详细地讨论这些情况，即从下面开始来考察GGC及其关系$\mathscr{R}$。

如前面所述，基本观点是：遗传密码是受条件控制的关系；因此，F_1和F_2应该当作关系来处理；并认为前面推断出的对称性基本假定将适用于在合适的生物环境条件下的翻译密码的生命过程。这种对称性的拓广是因取消了映照的特征所致，这样就可以假定某些关系是密码子简并性的必要(但不是充分的)条件。对于 2 级简并，关系定义如下：

$$R_1=\{(d_1^{(a,2)},d_2^{(a,2)})\in\mathscr{C}\times\mathscr{D}\mid d_1^{(a,2)}\in\mathscr{C}_U,d_2^{(a,2)}\in\mathscr{D}_C\}, \tag{7.86a}$$

$$R_2=\{(d_1^{(a,2)},d_2^{(a,2)})\in\mathscr{C}\times\mathscr{C}\mid d_1^{(a,2)}\in\mathscr{D}_A,d_2^{(a,2)}\in\mathscr{D}_G\}. \tag{7.86b}$$

---

* 关于同素异型式的数学处理请参考：

Findley, G. L.(1982)Int. J. Quant. Chem., Quant. Biol. Symp., 9:59—63.

表7-8 $\mathscr{C}$的群理论分解

记号	$\mathscr{C}$ 的子集			
	C_U	C_C	C_A	C_G
D_1	CAG	UAG	GUC	AUC
	UCC	CUU	AGG	GAA
	UAA	CGG	AUU	GCC
	UGG	CAA	ACC	GUU
	UUU	CCC	AAA	GGG
D_2	UAG	CAG	GUC	AUC
	UCC	CUU	GAA	AGG
	CGG	UAA	AUU	GCC
	CAA	UGG	ACC	GUU
	UUU	CCC	GGG	AAA
D_3	CAG	UAG	AUC	GUC
	CUU	UCC	AGG	GAA
	UAA	CGG	GCC	AUU
	UGG	CAA	GUU	ACC
	CCC	UUU	AAA	GGG
D_4	AUC	UAG	GUC	CAG
	GAA	CUU	AGG	UCC
	GCC	CGG	AUU	UAA
	UGG	ACC	CAA	GUU
	UUU	AAA	CCC	GGG
D_5	CAG	GUC	UAG	AUC
	UCC	AGG	CUU	GAA
	UAA	AUU	CGG	GCC
	GUU	CAA	ACC	UGG
	GGG	CCC	AAA	UUU
D_6	GUC	UAG	CAG	AUC
	AGG	CUU	UCC	GAA
	UAA	GCC	AUU	CGG
	ACC	CAA	UGG	GUU
	UUU	GGG	AAA	CCC
D_7	CAG	AUC	GUC	UAG
	UCC	GAA	AGG	CUU
	AUU	CGG	UAA	GCC
	UGG	GUU	ACC	CAA
	AAA	CCC	UUU	GGG

每一密码子代表了其三个碱基的所有排列，例：CAG代表了CAG、AGC、GCA、GAC、和 ACG。

这个对于2级简并密码子集的假定的含义是：如果一密码子对是二度简并的，则其必须是R_1或R_2的一个元素。但有些密码子对虽是R_1或R_2的元素却不是二度简并的；因此上述条件是必要的(但非充分的)条件。对于4级简并密码子集，要四个密码子是四度简并的条件是：二个密码子组成一对属于R_1的元素之一的密码子对并且另外二个密码子组成一对属于R_2的元素之一的密码子对。同样，对于6级简并密码子集，六个密码子要是六度简并的必要条件是：二个密码子组成一对属于R_1的元素之一的密码子对并且另二个密码子组成一对属于R_2的元素之一的密码子对；剩下的最后二个密码子必须组成一对属于R_1或R_2的元素之一的密码子对。

用了R_1、R_2后意味着要选择一个特殊的集分解。实际上有八个关系集，每一个对应于D_0—D_7。有一点要注意：虽然集R_1和R_2由所有的分解D_0—D_7形成，但并不代表所有的$(64)^2$个可能的密码子对。这一定要这样，否则推广了的遗传密码将无环境条件可言。也就是讲，如果所有密码子对都能是某个R_1或R_2的一个元素，则必要性将失去意义。属于F_1或F_2元素的所有密码子对，对于任何分解D_0—D_3，也将是R_1或R_2的元素；即在映照f下的密码子简并度的对称性可用上述"关系"来规定。当用分解D_4—D_7时，那些密码子对并不是R_1或R_2的元素。实际上D_4—D_7对应的是多义密码子翻译方式。

必要性假定是从SGC对称性的基本形式引伸出来的。Findley假定以R_1和R_2来表示的这种对称性的一般形式可以满足遗传密码，而与生物环境条件无关；这一对称形式即作为GGC的关系的约束条件。例，一个二重简并密码子对应该至少是分解D_0—D_7其中之一的R_1或R_2里的一个元素。而分解D_4—D_7则隐含着异于SGC环境条件的生物环境条件的存在。

现在我们就用上面得到的一套方法来讨论多义性编码的情况。在表7-6里的所有离体条件下的密码子多义性翻译方式都可用关系R_1、R_2来给予合理解释。先考虑表7-6中的对应于只有二个密码子的氨基酸，每一个密码子对在一定的分解下满足2级简并必要性假定，如表7-9。因此可以说这些密码子对在某些适当的生物环境条件下是二度简并的。从表7-9还可以发现某些密码子对在D_4、D_5或D_6、D_7时满足2级简并必要性假定，但在D_1—D_3时不满足。这个情况证实了这一观点即D_4-D_7对于在非SGC的生物环境下的密码子简并度可能是适合的。

表 7-9 满足必要性条件的二度简并的离体多义性编码

氨基酸	密码子对	分解
Glu	(UGA, GGA)	D_6, D_7
Gln	(CUA, UGA)	D_4, D_5
His	(UCA, CCA)	D_1—D_6
Thr	(CAC, AAC)	D_6, D_7
Val	(UGU, CGU)	D_1—D_6

表7-6里的其他编码也有类似的情况：

① 对应于Asp的六个密码子在任何分解下不满足6级简并必要性假定，即可以说这些密码子不会是六度简并的。然而，可以看出在一个分解下其中四个是四度简并的，而另二个是在另一

分解下二度简并的。结果是可以认为在一套多义性密码子翻译方式之中可能存在不同的生物环境条件。

② Tyr的情况类似于Asp。六个Tyr密码子不能成为六度简并的，但可以是一套四度简并密码子和一个二度简并的密码子对。

③ 对应于 Arg 的四个密码子不能 成为四度简并的，但可以 有 好几 种二 度简并的密码子对。

④ Cys的情况类似于Arg。四个Cys密码子不是四度简并的，但可以有二个二度简并密码子对。

⑤ Lys的情况相当特别，CUA 在不同的分解下和 AAA 或 AAG 可构成合条件的一对密码子，而后两者是在f映照下对应于 Lys 的密码子。因此，似乎是CUA翻译成Lys的方式是一种略作修正的 f 映照。

⑥ 对应于Gly有三个密码子，如果把GGG也加进去的话，则这四个密码子在 D_6、D_7 分解下满足4级必要性假定。事实上，GGG是在f映照下对应于G/y的密码子，因此这四个密 码 子 可认为是四度简并的。

⑦ 最后来看表7-6中非简并的密码子。如结合f映照，多一个上述密码子都可以在f映照里找到一个相应的密码子而成为一对在某一分解下符合 2 级必要性假定的密码子对。 因此，在这些情况下，这些密码子对可以认为是二度简并的。

再来看体内的情况，上面已提及有抑制基因突变时会产生体内的编码多义性。仅举三 个 例子，来说明必要性假定也适用于体内的多义性编码。

① Hirsh[40]曾研究了一个无义的大肠杆菌的抑制基因，其中UGA对应于 Trp。 在 D_0—D_5 分解下，UGA和f映照下的Trp密码子UGG满足 2 级必要性假定。因此，这二个密码子可认为是二度简并的。

② Berger和Yanofsky[41]发现一误义的抑制基因，其中Asp密码子被翻译成Gly。从 f映照里，所有对应于Asp的密码子和上述Gly密码子可在一定分解下形成一对满足 2 级必要性假定的密码子对。但GAU、GGG、GAC、GGG例外。

③ Carbon[42]等发现一误义抑制基因，其中一Arg密码子被翻译成Gly。和上面一样，在 f 映照里的Arg密码子可和Gly密码子形成二对满足 2 级必要性假定的密码子对，即 AGG、GGG和AGA、GGA。

近年来，通过对线粒体 DNA(mtDNA) 的碱基顺序分析及其相应的蛋白质氨基酸顺序的测定，发现其密码翻译方式有不少例外[5,15]，如表7-10；而利用 GGC 理论，这些例外 都 可以得到解释[43,44]；在不同的分解下，表7-10里的每一个 mtDNA 对应的密码子都可以找到一个在 SGC 里与之等价的密码子而组成一对符合GGC对称性要求的密码子对。表7-10，也清楚地表示了 在线粒体和胞质蛋白质合成时，生物环境条件变化时密码子翻译方式的不同。

上面，我们详细介绍了广谱遗传密码理论。现在，我们可以来总结一下 Findley 的这一密码子理论的思路：根据Gatlin的观点，认为生物环境条件的不同是密码子多义性的根源。这是GGC理论的核心。从SGC的对称性，归纳出密码子简并的对称性要求，并找出在映照概念下的对称性

表 7-10 线粒体不同编码方式的对称性

各种方式	等价编码*	分解	种
UGA→Trp	UGG	$D_0—D_8$	人类、牛、鼠、 果蝇、酵母属、 曲霉属、链孢霉属
AUA→Met	AUG	$D_0—D_3$, $D_6—D_7$	人类、牛、鼠、 果蝇、酵母属
AGA→Ser	UCU UCC UCA UCG AGU AGC	$D_4—D_7$ $D_6—D_7$ $D_2—D_3$ $D_0—D_1$ $D_4—D_6$ $D_6—D_7$	果蝇
CUA→Thr	ACC ACG	$D_1—D_3$ $D_0,D_6—D_7$	酵母属
CUG→Thr	ACU ACA ACG	$D_1—D_3$ $D_0,D_6—D_7$ $D_4—D_5$	酵母属
CUC→Thr	ACU ACA ACG	$D_0,D_6—D_7$ $D_1,D_4—D_6$ $D_2—D_3$	酵母属
CUU→Thr	ACU ACC ACA ACG	$D_4—D_5$ $D_0,D_6—D_7$ $D_2—D_3$ D_1	酵母属
AGA→TC	UGA UAG	$D_4—D_5$ $D_0,D_4—D_6$	人类、牛
AGG→TC	UGA UAA UAG	$D_0,D_6—D_7$ $D_0,D_4—D_5$ $D_6—D_7$	人类、牛
AUA→起始	AUG AUU	$D_0—D_3$, $D_6—D_7$ $D_4—D_5$	人类、牛、鼠 果蝇
AUU→起始	AUC AUA	$D_0—D_3$, $D_6—D_7$ $D_4—D_5$	人类、鼠 果蝇
AUC→起始	AUG AUU	$D_4—D_6$ $D_0—D_3$, $D_6—D_7$	鼠

* 指SGC码，起始密码子除外。

条件F_1、F_2；再把这一套方法作引伸和类比，建立GGC理论及其密码子简并的对称性要求，并在关系概念下作出相应于SGC的对称性条件假定R_1、R_2。（这里要提一下，GGC的对称性条件比SGC的条件弱。）从而把GGC理论应用于各种情况下的密码子多义性现象，几乎所有的多义性密码子都符合GGC理论的规则，Findley认为这一结果支持了GGC理论和密码子多义性是遗传密码整体的一部分的观点。

但要指出：Findley的理论，本质上是一种遗传密码的形式理论，虽然密码子的多义性通过密码子简并的对称要求从侧面在形式上给予了分析，但密码子多义性的现象仍没有从根本上得到解决。此外，该理论的假定条件较复杂、较弱，物理内容不够，似乎不是很有说服力。不过，Findley的观点和思路或许是正确的。量子遗传学期待着更理想的遗传密码理论的出现。

3）关于密码子的专一性计算　我们已经知道mRNA翻译成蛋白质时，并不是在模板mRNA上直接进行的，而是通过tRNA携带氨基酸去识别密码子等一系列专一的酶参与的复杂过程。但是在遗传密码最初的决定阶段，很难想象就已有高度专一性的酶参与使tRNA与氨基酸结合，又根据在实验中也发现多肽和碱基、聚精氨酸、聚赖氨酸和聚核苷酸之间有专一的相互作用[45-47]；因此有些意见认为在开始阶段，tRNA碱基和氨基酸直接的相互作用是决定密码的基础。Rendall[48,49]等按照这种观点，用扩展的Hückel方法（extended Hückel molecular orbital,EHMO）和微扰法计算了堆积的碱基和甘氨酸的相互作用，见表7-11，相互作用能包括静电、极化、色散三部分。碱基和甘氨酸的相对位置如图7-30。结果发现，能量值的大小和碱基的种类有密切的关系。甘氨酸的密码子是：GGU、GGC、GGA、GGG；因此，其相应的tRNA上的反密码子是CCA、CCG、CCU、CCC，简并了的反密码子是CC。从表7-11可知，CC堆积和甘氨酸具有很大的相互作用能量。因而，甘氨酸的三联体反密码子的专一性很可能是因其直接相互作用决定的。Rendall的计算只是一种初步的尝试。

4）关于密码子简并性的计算　Grick曾用变位假说来解释密码子的简并性。永田亲义[50]等从碱基及其互变异构体的稳定性计算来考虑密码子的简并性。从表7-12可以看出，肌苷、鸟嘌呤的异构体能量低，较稳定。因此，可以推测肌苷、鸟嘌呤混有相当量的异构体。根据氢键位置的不同，鸟嘌呤、肌苷的异构体和尿嘧啶配对，尿嘧啶的异构体和鸟嘌呤、尿嘧啶配对，结果简并性如表7-13所示。这一结果和Grick的从变位假说推出的简并性结果完全一样。前者是从能量来考虑的，后者是从主体条件来考虑的。

7.2.2.3 转译模型

在核糖体上进行蛋白质合成的基本物理问题是研究密码子反密码子识别的机制，肽键形成机制及核糖体沿着mRNA移动的机制。

密码子识别反密码子的过程极为精确，其误差只有万分之几，它们精确的配对不只是由热力学因素即各种配对之间的能量差别决定。Ninio[51]对密码子和反密码子三联体的水平和垂直相互作用进行了量子力学计算后，发现识别过程不是能量上唯一的。这种非唯一的识别也有“变位”效应(wobbles)的因素。错误的密码-反密码子对和正常的密码-反密码子对的相互作用能量差别不多；GUG和CAC、GUG和CCC的相互作用能量仅差8.8kJ/mol。因此，为迴避这一困难，Ninio曾提出“剩余三联体”(superfluous triplets)假说，即引起非唯一性识别的反密码子并

表 7-11 用推广的Hückel法所得碱基的重迭和甘氨酸相互作用的能量（kcal/mol）

碱基重迭*	静电相互作用能	色散能	极化能	合计
GC	45.5	22.2	25.2	92.9
GG	51.3	14.2	19.4	84.9
CC	45.2	16.0	21.7	82.9
CA	35.1	19.9	23.8	78.8
AC	34.9	21.4	21.0	77.3
UG	39.3	11.7	15.9	66.9
GU	38.3	9.4	13.5	61.7
AA	27.6	12.9	15.8	56.3
CG	29.0	9.8	14.6	53.4
CU	25.5	10.9	12.3	48.7
UA	20.8	14.1	13.1	48.0
UC	25.5	8.9	12.0	46.4
GA	25.3	8.7	10.1	44.1
UU	23.6	8.4	10.7	42.7
AG	21.6	7.8	10.3	39.7
AU	19.3	6.5	9.5	35.3

* GC 表示C的重迭，假定 DNA 是 B 型配置。
G（C下方竖线连接G）

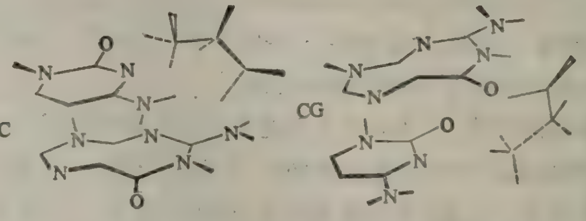


图 7-30 甘氨酸和GC及CG重迭的立体配置(对应于能量最低值)，甘氨酸在两碱基的重迭距离(0.34nm)的中间

不存在，当然这一假说和分子遗传学及分子生物物理学一般概念相矛盾。而且转译不是一平衡态，而是一有酶参与的动力学过程。

识别不仅是由热力学的因素决定，也由动力学因素决定：各种tRNA反密码子能以不同的比率和mRNA的密码子相互作用，这决定于tRNA和核糖体的结构和性质。

Woese[52]提出考虑核糖体性质的识别模型：识别不能简单地认为只是密码子-反密码子相互作用；密码子被tRNA和核糖体分别读二次。第一次认读只是大概的，核糖体并不区别密码子上的所有字母，但由于其变构性质，核糖体允许某些tRNA和密码子相互作用。这样，该模型提出几种类别的tRNA存在，这些类别可由核糖体来区别。不仅是反密码子结构，还有整个tRNA分子也起作用，tRNA的"关键"联系部位接近于反密码子区域。Woese把tRNA根据物理观点进行了分类；密码子有使核糖体构象变化的作用。实验表明，核糖体的状态强烈影响转译的精确性。

表 7-12 RNA 碱基的正常型(酮,氨基结构)和异构体(烯醇,亚氨基结构)的总能量(用CNDO/2法)

碱 基	正 常 型	异 构 体	能 量 差 (eV)
I	−2817.14	−2817.76	0.62
G	−3156.11	−3156.68	0.57
U	−2489.58	−2489.13	−0.45
C	−2326.68	−2325.53	−1.15
A	−2655.09	−2652.97	−2.12

* I表示肌苷，G，U，C，A 表示 4 个 RNA 碱基。

表 7-13 考虑了异构体结构对的 Watson-Crick 型氢键对的反密码子和密码子的碱基对

反密码子上的第一个碱基	密码子上的第三个碱基
U	A
U⁺	{G, U}
C	G
A	U
G⁺	U
G	C
I⁺	U
I	{C, A}

+ 符号表示异构体。

模型也假定tRNA有构象转变，当tRNA和核糖体相互作用时，并再和mRNA 作用时，就有这种构象转变，所以转译也是一个变构过程。但后来的实验表明，转译决定了 mRNA 的构象[53]，并涉及mRNA的降解。

singh[54]等提出转译的动力学过程理论。核糖体结合在mRNA的5′端，外切核酸酶也从5′端开始作用。核糖体保护了mRNA的端部使其免受降解作用。当核糖体移位时，5′端或降解或结合一新的核糖体。降解了的一端不能结合核糖体，而已结合了的核糖体则向前移位继续进行蛋白质合成。对于一随机过程，在时间t内，5′端存在的机率是$\exp(-kt)$，在聚核糖体 P_1、P_2、$\cdots P_{n-1}$上(这些聚核糖体分别含有1、2、$\cdots n$-1个核糖体)，在时间$(n-)t$内，合成的蛋白质分子数等于：

$$(1-e^{-k})+2e^{-k}(1-e^{-k})+\cdots+(n-1)e^{-(n-2)k}(1-e^{-k}) \tag{7.87}$$

在这一表达式中时间用单位t计算。当$t\to\infty$时，聚核糖体P_n上所合成的蛋白质分子数是：

$$e^{-(n-1)k}+e^{-nk}+\cdots+(n-1)(1-e^{-k})(e^{-(n-1)k}+e^{-nk}+\cdots) \tag{7.88}$$

将式(7.87) 和式(7.88) 相加，则得在无限时间内在一个 mRNA 链上合成的蛋白质分子总

数：

$$N \doteq (1-e^{-k})^{-1} \tag{7.89}$$

对于稳定的mRNA，$k=0$，$N\to\infty$；对极不稳定的mRNA，$k\to\infty$，$N\to 1$。可以根据观察到的标记链的衰变来得到mRNA降解的情况：可以认为：5′端的偶然失活，连着 的是一 个由核糖体移位决定其速率的降解步骤，造成了衰变。如果标记了一段长度为$(n-1)\delta$的 mRNA 链，则在时间$n-1$内，长度为δ的保持完整的单位数是

$$M_0=(1-e^{-k})[e^{-k}+2e^{-2k}+3e^{-3k}+\cdots+(n-2)e^{-(n-2)k}+(n-1)e^{-(n-1)k}]$$
$$=[e^{-k}(1-e^{-(n-1)k})]/(1-e^{-k}) \tag{7.90}$$

而在任何相继时刻t的完整的单位数是：

$$M_t=[e^{-(t+1)k}(1-e^{-(n+t-1)k})]/(1-e^{-k}) \tag{7.91}$$

而生存到时刻t的单位数f是

$$f=M_t/M_0=[e^{-kt}(1-e^{-(n+t-1)k})]/(1-e^{-(n-1)k})$$
$$n\gg 1,\quad f=e^{-kt} \tag{7.92}$$

真实的情况决定于转录和转译的相对速率。根据 Maaloe[55]的意见转录较快，但 Stent[56]认为这二步的速率数量级相同。而实验结果倾向于后一种意见。用 Singh 的理论 可以计算 mRNA 的长度分布，对标记的细菌 mRNA 作沉降分析，结果表明从 4 S到 80 S 有一放射性的宽带，其峰位置在 8S～12S范围内，这一结果和 Singh 的计算相似。但这一 理论 并没有 能解决核糖体等作用的详细机制。

7.3 量子遗传学现状

从量子遗传学近年来的发展看，大致有两种趋势。一种是紧跟着分子遗传学的发展，力图从定性上阐明分子遗传学出现的新现象、新问题，如 Z-DNA，线粒体密码子等；另一种是应用更完善的的量子化学从头计算程序从定量上更精确地计算 DNA 双螺旋的能量及 微 观 结 构 数 据，DNA 和水、离子环境的相互作用，DNA 和蛋白质的相互作用等，以更深入了解生命信 息 流的微观机制，下面主要从这两方面作些介绍。

自 1979 年 Rich[57]等报导了存在左旋 DNA 双螺旋(Z-DNA)结构以后， 引起了分子生物学界的很大震动。之后，曾发现在有些非常接近于生物体系的条件下，左旋 DNA 双螺旋可以稳定存在，因此自然就产生了一个这样的问题：左旋 DNA 是否在生物体系中存在，如果存在的话，其作用是什么？Rich 等将 Z-DNA 注入实验动物后发现产生了一种对 Z-DNA 构 象 极专一的抗体，这种抗体和右旋 DNA 完全不发生作用。后来， Rich 等又以实验表明 Z-DNA 构象片段很可能存在于具有生物活性的染色体上。因此 Sarma[58]认为 Z-DNA 的生物作用很 象和 基因组表达的控制有关。但 Rich 用的是 d-CGCGCG 单晶，并认为高 盐状 态下 的聚(dG-dC)·聚(dG-dC)即这种结构，因此，首要的问题是应考察从该种低聚核 苷酸模型衍生出的聚(dG-dC)·聚(dG-dC)结构和溶液状态下的这种 DNA 双螺旋结构接近的程度。Sarma 等计算了各种型式的聚(dG-dC)·聚(dG-dC)结构的理论核磁共振屏蔽常数，并和高、低盐状态下 的聚(dG-dC)·

表 7-14　聚(dG-dC)·聚(dG-dC)的屏蔽常数的实验值和理论计算值

质子	模型	计算的(Δδ)	高盐 Δδ	低盐 Δδ
	A-DNA	1.49		
CH5	AH/B-DNA	0.85		
	Z-DNA	0.95	1.05	0.82
Levitt 螺旋	B-DNA	0.71		
Dickerson 螺旋	B-DNA	0.81		
	A-DNA	0.20		
CH6	AH/B-DNA	0.73		
	Z-DNA	0.51	0.62	0.70
Levitt 螺旋	B-DNA	0.34		
Dickerson 螺旋	B-DNA	0.36		
	A-DNA	0.45		
CH1′	AH/B-DNA	0.58		
	Z-DNA	0.91	0.69	0.56
Levitt 螺旋	B-DNA	0.19		
Dickerson 螺旋	B-DNA	0.28		
	A-DNA	−1.12		
GH8	AH/B-DNA	−0.15		
	Z-DNA	0.18	0.30	0.27
Levitt 螺旋	B-DNA	−0.44		
Dickerson 螺旋	B-DNA	−0.80		
	A-DNA	−0.21		
GH1′	AH/B-DNA	0.25		
	Z-DNA	0.17	0.15	0.11
Levitt 螺旋	B-DNA	−0.35		
Dickerson 螺旋	B-DNA	−0.21		

聚(dG-dC)实验数据进行比较(如表 7-14)。

发现,在五种不同的 DNA 空间构型的理论屏蔽常数中,在合理的精度范围内, 高盐状态下五个质子部位的实验值和 Z-DNA 的计算值全部符合,其中 CH_5、CH_6、GH_8、GH1′符合得很好(差距在 0.02～0.12 ppm),只有一个CH1′差距稍大些(0.22 ppm)。因此可以推测: 高盐型式的聚(dG-dC)·聚(dG-dC)是左旋的 DNA。这和 Rich 等用 Raman 光谱对这二种结构测定的结果一致。同时从表 7-14 也可以看出,低盐状态下的实验数据和经典的 Arnott 及 Hukins B-DNA 结构的理论值在CH_5、CH_6、CH1′三个质子部位符合得极好,CH1′差距稍大些(0.17 ppm),只有 CH_8不符;因此推测低盐状态下的聚(dG-dC)·聚(dG-dC) 结构从时间平均和涨落平均来看是接近于 AH/B 型式的,这和质粒中聚(dG-dC)·聚(dG-dC)克隆片段实验所表示的这种片段之 B 型乙型转变情况一致。Arnott[59]和Chandrasekaran、Olson[60]分别从理论上讨论了B→Z 转变中的空间结构变化。要提一下,当初曾通过晶体学研究发现了和传统 Watson-Crick对不同的碱基氢键配对形式,这些曾被斥之为实验室中的怪事;后来当 tRNA 晶体结构阐明之后,才发觉这些异常的碱基对确实存在。Z-DNA 的情况也有些类似,Z-结构是在离体条件下直接从

单晶结构观察到的。所以可设想：既然大自然经过长达 40 亿年的进化过程巧妙地造出了一种能右旋又能左旋的可变的分子结构，那么这种结构一定有它的生物功能。Z-DNA 的生物作用是什么？这也正是遗传学要去解决的问题。

遗传密码子的简并性问题，也是近年来量子遗传学的兴奋点之一。已知有 61 个密码子对应于 20 种氨基酸，如果平均分配的话，则每种氨基酸将有 3 个密码子与之对应；但实际上各种氨基酸所对应的密码子数目并不都是 3，即对应于某种氨基酸的密码子的简并度不全是 3，简并度有从 1 至 6 各种情况，如亮氨酸有 6 个同义密码子(synonymous codons)，而色氨酸只有一个密码子。并且还发现：对应于三个或三个以上同义密码子(下称"同义子"，synonyms)的氨基酸，它们的化学结构最简单，其特征是具有脂族侧链及羟基官能团；而对应于二个同义子或一个密码子的氨基酸，其残基则含有芳环、硫、酰胺或带电荷的官能团。通过分析蛋白质氨基酸组成又发现：一种氨基酸其对应的同义子越多，则其在蛋白质分子中出现的机会越大。如何来解释这些情况？为什么简并度不一样？Dufton[61]研究了这一饶有兴趣的问题。Dufton 认为密码子的简并度不一，可能和保证蛋白质二级结构的进化稳定性有关：根据密码子三联体随机突变及很大程度上是非倾向性突变(neutral mutation)的基本原则，可以说明上述情况。密码子简并度的模式的意义在于：通过非倾向性突变产生一种一般的进化趋势即倾向于使在蛋白质片段中化学上较简单的及结构上较不易改变的氨基酸不经受严格的选择，其效果是减少了蛋白质分子中支持结构部分的改变而使含有较复杂化学结构之氨基酸的功能部分有改变的可能。这样就有利于进化。Dufton 将蛋白质中的氨基酸实际组成和理论组成(即：完全由同义子出现的几率决定的组成)进行了比较；还将蛋白质组成数据按二级结构进行了分类，则使得同义子的分布情况和多肽构象的不同类型相联系，如表 7-15。由于具有较多同义子的氨基酸其出现的机会较多并且其结构又简单不易变化，因此高、低同义子氨基酸含量上的差别将影响二级结构进化的倾向性和稳定性。Dufton 还计算了在每一种二级结构里的每个氨基酸的平均同义子数，如表 7-16。从表中可以看出，这四种二级结构的平均同义子数接近(3.52～3.36)，这表明每种结构都会有以和三个以上同义子对应的氨基酸为主的成份；因此虽然每个氨基酸形成各种二级结构的倾向不一，但每种二级结构对氨基酸的平均来说并无差异。由此看来，同义子采取这种分布的意义在于：蛋白质中的氨基酸随着进化过程可以由自然选择而改变，但蛋白质的构象是保持稳定的。从已研究过的许多蛋白质谱系来看，都表明了它们的基本构象保留着，而其功能及专一性方面却有较大的变化。

近些年来，对于 DNA 和环境的相互作用的量子化学计算进展较大[62]，对于以整个核苷酸作为元胞的能带结构计算也取得了实质性的进展[63]，关于水结构对于核苷酸碱基对的影响[64]，聚核苷酸-聚多肽的相互作用[65]、DNA 大分子的非周期性处理[68]等等的计算，都进行了探索。DNA-蛋白质相互作用决定了基因的表达，而对聚核苷酸-聚多肽之间的相互作用能量的计算，则是朝着解决这一问题的第一步。Ladik[65]等近来曾用从头计算法对聚甘氨酸链分别和同聚核苷酸：聚(TSP)或聚(ASP)(这里：S 表示糖、P 表示磷酸酯、T 代表碱基)的相互作用能进行了大规模的计算。这些计算中，对聚甘氨酸链的几种构象包括全伸展的、2-螺旋的分别作了能带结构计算；但当周期性的同聚核苷酸和周期性的聚甘氨酸各处于对方的场中时，则变得非周期性

表 7-15 与遗传密码相关的多肽二级结构分布①[66]

氨基酸②	综合数据	α-螺旋	β-折叠	β-回折	随机线圈
Lys(2)	2.13	2.62	1.65	2.04	2.16
Asp(2)	1.83	1.90	1.31	2.58	1.68
Glu(2)	1.55	2.24	1.16	1.55	1.10
Gly(4)	1.43	0.80	1.32	2.35	1.65
Asn(2)	1.34	1.21	1.00	1.71	1.62
Ala(4)	1.28	1.65	1.16	0.99	1.19
Val(4)	1.22	1.10	1.81	0.57	1.22
Tyr(2)	1.16	0.85	1.47	1.22	1.16
Ile(3)	1.10	1.07	1.59	0.56	0.96
Phe(2)	1.07	1.16	1.41	0.63	0.91
Thr(4)	0.98	0.80	1.18	1.01	0.91
Gln(2)	0.94	1.19	0.76	0.92	0.85
Trp(1)	0.91	0.92	1.06	0.71	0.98
Met(1)	0.91	1.39	0.90	0.38	0.79
Ser(6)	0.81	0.66	0.77	1.07	0.82
His(2)	0.73	0.89	0.78	0.49	0.64
Leu(6)	0.70	0.90	0.71	0.41	0.65
Cys(2)	0.67	0.74	0.49	0.54	0.91
Pro(4)	0.61	0.32	0.38	1.16	0.81
Arg(6)	0.30	0.29	0.31	0.27	0.38

① 观察值%/根据同义子分布预期的值%
② 括号内数字是同义子数。

表 7-16 每个氨基酸的平均同义子

类型	每个氨基酸的平均同义子	每个五肽的编码
61个密码子内取样	3.85	846
β-折叠	3.52	540
随机线圈	3.51	533
β-回折	3.50	525
整个数据集	3.46	496
α-螺旋	3.36	428
20种氨基酸内取样	3.05	264

了，因此需用互洽场方法(mutually consistent field (MCF) method)。交换相互作用则用微扰理论来解决，而分散相互作用仍用 London 的式子。计算结果表明：对全伸展或 2-螺旋构象的聚甘氨酸和聚(TSP)或聚(ASP)的 Hartree-Fock 水平相互作用能为几个负 kcal/mol/核苷酸单位、甘氨酸单位。而分散能的修正值只有负几分之一 kcal/mol，这说明它们的相互作用基本是静电性质的。

Hosur[67]等计算了肽键骨架和核酸碱基对模型的相互作用。此相互作用模型是肽键骨架在碱基对间形成氢键桥。肽键的方向和碱基对位置有两种情况，如下图 7-31。因此在肽键和碱基

对通过二对氢键连接后形成一环，一种情况是八原子环，另一种情况是九原子环，两种情况的肽键骨架的相对构象是有明显区别的，这可能和专一性识别的原因有关[66]。这两种情况的相互作用能量也不同，如表 7-17，九原子环的似乎更稳定些。从表 7-17，还可以看到：聚(L-Ser)和 A-U 对及 G-C 对的相互作用能差别较大，如果这种差别对其它肽键说来也如此的话，那么这将可能是蛋白质识别碱基的主要因素了。Hosur[67]等还对聚(L-Ser)和聚(A)·聚(U)的复合物测定了 ^{13}C NMR 光谱，实验表明，聚(L-Ser)和聚(A)·聚(U)双螺旋以二对氢键连接，和其设想的模型一致。

表 7-17　碱基对和三丝氨酸之间的稳定化能①

肽的构象	碱基对	环的大小	稳定化能 (kcal/mol)
A	A-U	8	−0.7
B	G-C	8	−6.2
C	A-U	9	−6.5
D	A-U	9	−2.7
E	A-U	9	−0.1
F	G-C	9	−10.1
G	G-C	9	−5.0

① 稳定化能等于相互作用能减去构象能。

Pullman[68-72]等对酵母 $tRNA^{Phe}$的分子静电势、原子的立体可接近度(molecular electrostatic potential、atomic steric accessibility)和环境中水分子及金属离子对其的影响作了一系列量子化学计算。计算结果可用"势轮"和"接近度轮"形象地表示，如图7-32和图7-33。在"势轮"中，圆周上的数字表示碱基编号，1至76；圆的半径表示势的尺度，−600至−1350 kcal/mol。每一碱基的势用一径向线段表示，线段两端分别是最大的和最小的势。从图可以看出，结构和势的相关很强，在 $tRNA^{Phe}$各个环及受体端部位的势明显较弱，而螺旋结构的势较强，在靠近分子中心处最强。金属离子对 $tRNA^{Phe}$的静电势屏蔽效应较大，但"势轮"的基本形状不变。图 7-32 是 $4Mg^{2+}$屏蔽了的 $tRNA^{Phe}$的"势轮"。在"接近度轮"中，圆周上的数字仍表示碱基的编号；半径表示可接近的面积，0 至 15 i^2，圆心处可接近度最大。每一碱基的可接近度用一径向线段表示，线段二端的值分别表示该碱基部位之最大和最小的可接近度。从图也可以看出，可接近度和结构有较强的相关关系。$tRNA^{Phe}$各环和受体的碱基之可接近度大，而其螺旋结构部位之碱基的可接近度小。静电势是以磷酸酯的长程效应为主的，而可接近度是以短程空间障碍为主的，对局部结果敏感。因此静电势的结构相关要强些，而可接近度对 $tRNA^{Phe}$ 的整体结构相关不如静电势的明显，这从比较两种"轮子"可以看出来。这些"轮子"可以用以讨论 tRNA 的反应性。关于可接近度的计算方法可参考 Lee[73]和 Richards 的工作。

Pullman[74-81]等还计算了 B-DNA 螺旋的分子静电势和金属离子对这螺旋的屏蔽效应，得到了双螺旋内部大沟道和小沟道的各种势深度的分布。大、小沟道对于专一性相互作用都起作用。Pullman 的这一理论计算结果可用于解释乳糖阻遏物和 DNA相互作用的实验结果[82]，此实验结果是阻遏体非专一地和非操纵子 DNA 的小沟道结合，而阻遏体和操纵子DNA的两种沟道

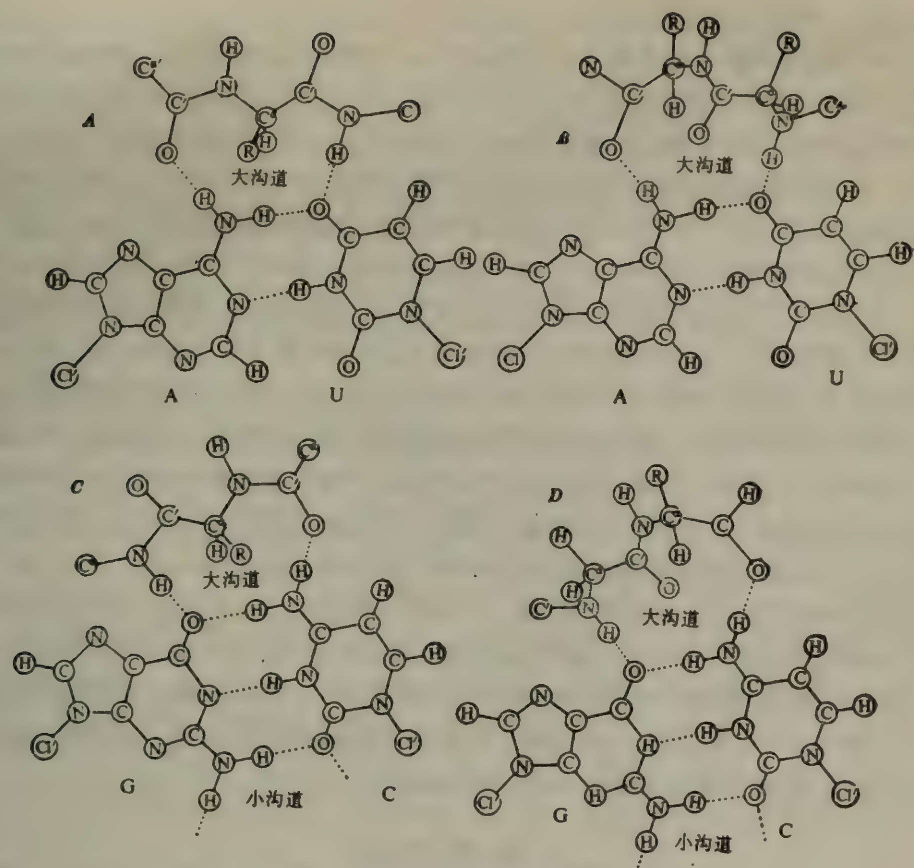


图 7-31 肽键骨架和碱基对相互作用示意图，肽键骨架在氢键桥中有两个相反方向，分别形成八原子环（A、C）和九原子环（B、D）。



图 7-32 tRNA[Phe] 碱基的"势轮"

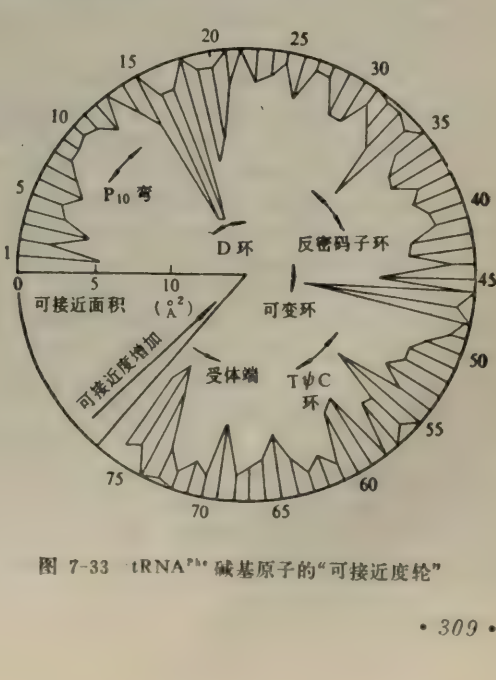


图 7-33 tRNA[Phe] 碱基原子的"可接近度轮"

的结合都是专一性的相互作用。

DNA 复制的精确性也始终是量子遗传学中的一个重要问题。长久以来碱基互补对原则一直被认为是核酸复制的正确性的基础。1959 年以后，发现有非互补的氢键对存在的可能，因此根据互补对原则的DNA 复制的这种解释就发生了疑问，Loeb[83]等认为这是一种"酶控制"的过程，Löwdin[3]也早就提出过复制平面上涉及四个碱基的模型。Rein 等对在 DNA 合成过程中碱基选择时起重要作用的分子间相互作用及构象特征等一系列物理因素作了系统的研究[84-88]，比较了正常碱基对和错误配对时的各种物理因素的变化。Rein 认为复制过程中的生成自由能由下列因素控制：1）新配对的氢键能量；2）碱基对的堆积能；3）糖磷酸酯骨架形成时的构象调正所需的能量。前二个因素是分子间相互作用，由分子量子力学方法计算，计算中 Rein 等考虑了静电极化、分散和重叠排斥作用。还考虑了溶剂效应。第三个因素主要是熵的贡献，即由开启状态到闭合螺旋状态时，旋转自由度受限制所引起。Rein 等用由二对碱基对组成的最小螺旋模型（minihelix）计算了上述三部分自由能，并比较了这三种因素对复制正确性的贡献。结果表明：从生成自由能来看，有利于互补碱基对形成。互补对和非互补对的生成自由能差别至少在 29 kJ/mol 以上；碱基对配对的专一性主要是构象调正能量上有差别。

量子遗传学就作为一个专门理论的意义上来讲，刚刚起步，远未成熟。然而它是量子化学和分子遗传学覆盖而成的土壤上的一株可贵的嫩芽；分子遗传学是量子遗传学发展的基础，我们可以看到，量子遗传学的研究正在紧跟分子遗传学发展的步伐；任何以物质为基础的科学终究会深入到原子和电子水平。可以相信，最终解开遗传学之迷的希望还在于量子遗传学；随着量子遗传学的发展、成长、完善、遗传学的奥秘也将一步步地揭开。

参考文献

[1] Gardner, E. J. et al.: "Principles of Genetics", (John Wiley & Sons, New York), (1981).
[2] Watson, J. D. H.: "Molecular Biology of the Gene", (Benjamin, New York), (1965).
[3] Löwdin, P. O.: "Advances in Quantum Chemistry", Vol. 2, P. O. Löwdin (ed) (Acadmic Press, New York), (1965).
[4] 卢惠霖：《中国医学百科全书，医学遗传学》，上海科学技术出版社，(1984).
[5] 刘培南等：《基础分子生物学》，高等教育出版社，(1983).
[6] Delbrück, M. et al.: "Nachr. Biol. Ges.", (Wiss. Gottingen), 1: 189 (1935).
[7] Schrödinger, E.: "What is Life?", (Cambridge University Press, London), (1945).
[8] Gamow, J.: Nature, 173:318 (1954).
[9] Wiener, N.: "Cybernetics or Control and Communicati on in the Animal and the Machine", 2nd ed., (MIT Press and Wiley, New York), (1961).
[10] Löwdin, P. O.: U. S. Dep. Com. AD., 296941, 86pp, (1962).
[11] Löwdin, P. O.: Inter. Sci. Tech., (New York), (1963).
[12] Löwdin, P. O.: Rev. Med. Phys., 35:724 (1963).
[13] Löwdin, P. O.: Biopolymers Symp., 1:161 (1964).
[14] 王身立：自然杂志，3:886 (1980).
[15] Smith, E. L. et al.: "Principles of Biochemistry: General Aspects", 7th ed, (McGraw-Hill, New York), (1983).
[16] Volkenstein, M. V.: "Molecular Biophysics", (Academic Press, New York), (1977).
[17] 藤永憲："ガんの分子生物学"，讲谈社，(1976).
[18] Seybold, P. G.: Inter. J. Quantum Chem. QBS., 3:39 (1976).

[19] Delbrück, M. et al.: "The Chemical Basis of Heredity", W. D. McElroy and B. Glass (ed), (John Hopkins Press, Baltimore, Maryland), (1957).
[20] Meselson, M.: Proc. Natl. Acad. Sci. U. S. A., 44:671 (1958).
[21] Taylor, J. H. et al.: Proc. Natl. Acad. Sci. U. S. A., 43:122 (1957).
[22] Delbrück, M.: Proc. Natl. Acad. Sci. U. S. A., 40:783 (1955).
[23] Kornberg, A.: Rev. Mod. Phys., 31:200 (1959).
[24] Kornberg, A.: "Enzymatic Synthesis of DNA", (Wiley, New York), (1961).
[25] Cairns, J.: J. Mol. Biol., 6:208 (1963).
[26] Rich, A.: Rev. Mod. Phys., 31:50 (1959).
[27] Okazaki, R. et al.: Cold Spring Harbor Symp. Quant. Biol., 33:129 (1968).
[28] Pullman, A. et al.: "Quantum Biochemistry", (Wiley, New York), (1963).
[29] Drake, J. W. et al.: Genetics, 55:738 (1967).
[30] Baltz, R. H. et al.: Proc. Natl. Acad. Sci. U. S. A., 73:1269 (1976).
[31] Bingham, P. M. et al.: ibid., 73:4159 (1976).
[32] Cooper, W. G.: Inter. J. Quant. Chem., QBS., 6:171 (1979).
[33] Cooper, W. G.: Inter. J. Quant. Chem., 14:71 (1978).
[34] Rüger, W.: Biochem. Biophys. Acta, 238:202 (1971).
[35] Florentiev, V. et al.: Nature, 228:519 (1970).
[36] Geiduschek, E.: "Molecular Assoeiations in Biology", B. Pullman (ed), (Acad. Press, New York), (1968).
[37] Riley, P.: Nature, 228:522 (1970).
[38] Bibilashvili, R. et al.: Biochem. Biophys. Acta, 259:35 (1972).
[39] Yčas, M.: "The Biological Code", (North-Holland, Amsterdam), (1969).
[40] Hirsh, D.: J. Mol. Biol., 58:439 (1971).
[41] Berger, H. et al.: Science, 156:394 (1967).
[42] Carbon, J. et al.: Cold Spring Harbor Symp. Quant., Biol 34:505 (1969).
[43] Gatlin, L. L.: "Information Theory and the Living System", (Columbia Univer. Press, New York), (1972).
[44] Findley, G. L. et al.: Inter. J. Quant. Chem., QBS., 6:313 (1979).
[45] Findley, A. M. et al.: Inter. J. Quant, Chem., QBS., 11:109 (1984).
[46] Yuki, A. et al.: Biochem. Biophys. Res. Comm., 36:657 (1969).
[47] Wagner, K. G. et al.: Biochemistry, 7:1771 (1968).
[48] Rendall, M. S.: J. Am. Chem. Soc., 94:4337 (1972).
[49] Rendall, M. S. et al.: Biopolymers, 10:2083 (1971).
[50] Fujita, H. et al.: J. Theor. Biol., 28:143 (1970).
[51] Ninio, J.: J. Mol. Biol., 19:548 (1971).
[52] Woese, S.: J. Theor. Biol., 26:83 (1970).
[53] Fukami, H. et al.: Proc. Natl. Acad. Sci. U. S. A., 68:570 (1971).
[54] Singh, U.: J. Theor. Biol., 25:444 (1969).
[55] Maaloe, O. et al.: "Control of Macromolecular Sythesis", O. Maaloe and N. Kjeldgaard (ed) (Benjamin, New York), (1966).
[56] Stent, G.: Mendel Gentennial Symp. Roy. Soc., London, (1965).
[57] Wang, A. H. J. et al.: Nature, 282:680 (1979).
[58] Sarma, R. H.: Inter. J. Quant. Chem., QBS., 9:65 (1982).
[59] Arnott, S. et al.: "Biomolecular Sterodynamics", R. H. Sarma (ed), (Adenine Press, New York), (1981).
[60] Olson, W. K.: "Biomolecular Sterodynamics", R. H. Sarma (ed), (Adenine Press, New York), (1981).
[61] Dufton, M. J.: J. Theor. Biol., 116:343 (1985).
[62] Clementi, E.: "Structure and Dynamics of Nucleic Acids and Protein", E. Clementi and H.

Sharma (ed), (Acad. Press, Philadelphia), (1983).
[63] Ladik, J. et al.: Inter. J. Quant. Chem., QBS., 7:181 (1980).
[64] Otto, P. et al.: J. Chem. Phys., 77:5026 (1982).
[65] Ladik, J. et al.: Inter. J. Quant. Chem., QBS.,10:73 (1983).
[66] Day, R. S. et al.: Inter. J. Quant. Chem., 21:917 (1982).
[67] Hosur, R. V. et al.: Inter. J. Quant. Chem., 20:23 (1981).
[68] Lavery, R. et al.: Nucleic Acids Res., 8:1061 (1980).
[69] Lavery, R. et al.: J. Comp. Chem., 1:301 (1980).
[70] Lavery, R. et al.: Theor. Chim. Acta, 57:233 (1980).
[71] Lavery, R. et al.: Nucleic Acids Res., 8:5095 (1980).
[72] Lavery, R. et al.: Inter. J. Quant. Chem., QBS., 8:171 (1981).
[73] Lee, B. et al.: J. Mol. Biol., 55:379 (1971).
[74] Perahia, D. et al.: Theor. Chim. Acta, 50:351 (1979).
[75] Perahia, D. et al.: ibid., 51:349 (1979).
[76] Perahia, D. et al.: Inter. J. Quant. Chem., QBS., 6:353 (1979).
[77] Singer, B.: Prog. Nucleic Acids Res. Mol. Biol., 15:219 (1975).
[78] Hoogsten, K.: Acta Cryst., 12:822 (1959).
[79] Raszka, M. et al.: PRNAS, 69:2025 (1972).
[80] Pullman, B. et al.: Proc. Natl. Acad. Sci. U. S. A., 55:905 (1966).
[81] Pullman, B. et al.: J. Mol. Biol., 22:373 (1966).
[82] Pullman, B. et al.: Compt. Rend. Acad. Sci. Paris, 236:2006 (1966).
[83] Loeb, L. A. et al.: Cancer Res., 34:2311 (1974).
[84] Rein, R. et al.: Inter. J. Quant. Chem., 2:129 (1968).
[85] Claverie, P.: "Molecular Associations in Biology", Paris Symp., (1967).
[86] Egan, J. T. et al.: Inter. J. Quant. Chem., QBS., 1:71 (1974).
[87] Rein, R.: "Advances in Quantum Chemistry", Vol. 7, P. O. Lowdin (ed), (Acad. Press, New York), (1973).
[88] Rein, R. et al.: Inter. J. Quant. Chem., QBS., 2:145 (1975).

刘征先

第8章　量子药理学

8.1 引　言

量子药理学(quantum pharmacology)是一门用量子力学方法研究药理学问题的新兴科学。它是七十年代继化学药理学和分子药理学的发展以及量子力学中的分子轨道计算方法得到广泛应用而产生的边缘性学科。1977 年英国药理学家 Richards 编著出版了第一本介绍量子药理学的专著[1]。迄今，量子药理学作为一门新兴学科已被逐渐广泛地接受並获得了不断的发展。

随着科学的发展，药理学的研究逐渐从分子水平深入到了电子水平，从而使药理学学科本身从定性到定量的质的飞跃成为可能。我们知道，量子力学是描述电子运动的可靠数学定量方法。量子药理学的产生正是由于量子力学分子轨道方法在药理学中的应用，同时也由于药理学研究本身的迫切需要形成了不同学科的融合[2]。

早在 50 年代，药物化学家就尝试用简单分子轨道方法(HMO 法)对某些药物分子(如致幻剂)进行计算，开始了量子药理学探索。在六十年代以后，由于大型高速电子计算机的不断问世，各种分子轨道计算方法及相应程序的产生和改进，量子药理学的研究才获得迅速发展。

目前，分子轨道计算方法中近似最少、精确度最高的从头计算法，已经可以计算相当大的分子。进入 80 年代后，国外量子药理学研究的论文，约 1／3～1／2 系采用从头计算方法。虽然就目前计算的分子的大小而言，离药物化学家和药理学家的期望尚有一定的距离，况且完成从头计算所耗费的计算机机时之长又往往使人望而生畏。但是，可以预料，在不远的将来，从头计算将能胜任所有单个药物分子的电子信息的计算，甚至满足药物与碱基复合物计算的要求。现在作为一种过渡，使用半经验和经验的分子轨道法替代非经验的从头计算法，它们较易被药物化学家和药理学家所接受。因为他们不仅对单个分子感兴趣，而且特别对一系列同源化合物，进而对一大批分子的几何形状、构象乃至电荷分布的共同特征感兴趣。进行如此大量的分子轨道计算，使人们不得不借助于经验和半经验分子轨道计算方法。

令人欣慰的是，在这些带有很大近似的半经验分子轨道方法不断发展的同时，与药物研究息息相关的实验技术与方法获得了飞速的发展，如 X 射线结晶学；核磁共振(NMR)技术；偶极矩测定；高效液相色谱的广泛使用；以及计算机图形学(computer graphics)的诞生和发展。由这些实验所获的结果，如药物的空间三维结构与绝对构型，用 NMR 测得的溶液构象，分析偶极矩测量结果所得的电荷分布方向，以及用计算机图形学直观地显示出药物分子的三维图象等等，它们都可以用来检测量子化学半经验方法计算所推测的药物的物理性质，如构象、电荷密度、偶极矩等。与此相辅的是，用分子轨道近似方法计算的结果来讨论推测药物的理化性质，如油水分配系数 $\log P$，解离常数 $\mathrm{p}K_a$ 等也获得成功，这无疑又增强了药物学家对使用近似方法的信心。

诚然，这些理论计算一般是对孤立的气相分子在静止的物理状态下加以计算，而实验条件却

绝非如此，X 衍射是对晶体进行的，NMR 的测定是在相当浓度的溶液中进行，高效液相虽然可以在模拟生物体的酸碱性(pH=7.4)的条件下进行，但仍是一个很大的近似。这三种物理环境，没有一种恰巧符合药物分子在生物体内转运、代谢的真实的生理环境。

在六十年代中，由 Hansch 和 Fujita 等人开创並发展的药物定量构效关系（quantitative structure-activity relationships）研究恰好弥补了上述所谓难题。QSAR理论所建立的方程，即在同源化合物系列中导出众多物理化学参数的主要参数与药物的生物活性之间的相关方程。这些方程符合传统的统计理论並借助于计算机技术，这就为量子化学计算所得的众多电子信息，如能级、电荷密度、偶极矩、静电势等与该分子的药物生物活性之间建立了定量关系式，正由于QSAR研究成为药理学家与量子化学家之间的桥梁，使量子化学在药理学中的应用进人一个新的阶段。

从药理学研究的本身分析，把理论化学应用到药理学上来，除上述的可能性外还有其必然性。药理学中常把药物结构区分为非特异性和特异性两种类型。所谓非特异性药物的生物活性的产生，通常与药物的物理化学性质有关，例如油水分配系数、电离度、溶解度、表面活性等。所谓特异性药物，一般认为它们的作用先与生物体的大分子化合物或生物大分子组织结合成复合物，然后产生生化反应，继而显出不同的药理作用。由于药理研究的不断深入，受体分离和研究的新成果不断涌现及实验技术的不断提高，需要从分子-电子水平上了解和掌握药物-受体的三维电子结构分布，药物产生活性的药效构象——活性构象，这就必须应用量子化学计算。进入八十年代，越来越多的药理学家已感受到，这种需要决不是可有可无仅仅点缀而已。

量子化学理论计算带来的另一种优势是，它们不仅能够计算已经获得的药物分子，而且可以计算任何分子的性质，预测构象，不论对其整体或局部位置，都可方便地获得分子的各种信息，这些信息拿已经获得的活性结构相关的方程进行推测，便可预知未合成药物的可能活性或可能的药理作用机制。这无疑令人振奋和倍受欢迎。众所周知，按传统的药物设计方法是改造具有生物活性的先导化合物(lead compound)的结构中，采取“泛泛合成，随机筛选”的方法。据统计大约要合成 15 000个化合物方能找到一个临床使用的药物。且不论药物化学家在进行合成时所耗费的材料人力，光从药理筛选工作而言也是非常惊人的。当将量子化学计算与QSAR研究与药理学结合起来以后，一旦计算获得了对生物活性和药理作用机制的满意解释，就有可能不必首先合成化合物，在合成前，事先依据计算所获的信息，有目地有方向地改变影响活性的主要结构，即所谓“药物设计(drug design)”，可按设计后药物分子的结构进行量子化学计算，並预测具可能活性，然后进行统筹分析(主要取决于有机实验的难易)进行取舍，这将会大大减少合成和筛选的化合物数目。尽管“药物设计”获得完善成功的典型事例公布不多，然而，这无论如何是令人神往的，它将大大促进量子药理学的深入发展。

由于量子药理学是一门新兴学科，正处于迅速发展之中，亦由于它是边缘性学科，它所涉及的面较广，应该包括哪些内容尚在探讨发展之中。另外，由于对特定的药物的量子化学计算，包含了各种不同角度讨论药理现象，例如对组胺化合物，既有大量工作研究其构象，也有众多研究工作是探讨受体作用，而受体作用又与药物构象和电荷密度分子静电势等息息相关，不能截然分割。这种相互交叉给本章的通俗叙述带来了一定的困难。本章拟就药物分子的构象、药物的受体、药物分子的静电势和电荷密度、电荷转移、QSAR、药物动力学等几个大范围内介绍量子药理学

的基本内容。其中包括了药理学中常见的肾上腺素、多巴胺类、吗啡类、乙酰胆碱类、组胺类、抗癌药物、电荷转移复合物等药物的量子药理学研究现状及进展，使读者对前面所述的量子药理学发展的可能性和必要性有个较清晰的了解。

实际上，Richards 的“量子药理学”一书中 40%的内容介绍了分子药理学的知识，因此读者应对分子药理学内容有所掌握。我们认为，除参考“量子药理学”前半部分以外，可阅读 korolkovas 的“分子药理学概论”中译本[3]，还应参读我国药理专家撰写的《分子药理学》[4]、《神经药理学》[5]有关章节。若有的读者想了解量子药理学研究的历史，可参阅 Christoffersen 的综述[6]，该文将 1974 年前的量子药理学研究作了全面的阐述。而 1974 年至 1981 年间的研究论文可查阅 Richards第二版《量子药理学》中的文献目录[7]。所有这些对阅读本章内容将是非常有益的。

8.2 药物分子构象

8.2.1 构象对药物分子药效的重要性

研究证实，药物分子与生物大分子(DNA、核酸、酶、辅酶等)之间发生相互作用时，存在着空间化学问题。亦即这些生物大分子化合物有一定的构象(conformation)，且药物分子的构象只有与其匹配或适应，才易于与它结合，进而产生相应的药理作用[2]。从某种意义上讲，药物分子的构象是产生药效的决定性因素。(有关生物大分子的构象将在第 9 章讨论)

我们知道，组胺即 2-间二氮茂基乙胺，是人体各种器官及组织内广泛存在的化学介质。它作为一种神经递质，能引起体内多种药理效应。实验表明，体内至少存在两类组胺受体(receptor)——H_1受体和 H_2受体。组胺对这两种受体都有强烈的作用。如在组胺分子中引入 2-甲基或 4-甲基使之成为 2-甲基组胺和 4-甲基组胺化合物(图 8-1)后，由于甲基的引进，以及各自位置的不同，从而导致组胺分子构象变化的差异，这可能就是导致它们的药理效应差别的原因之一。

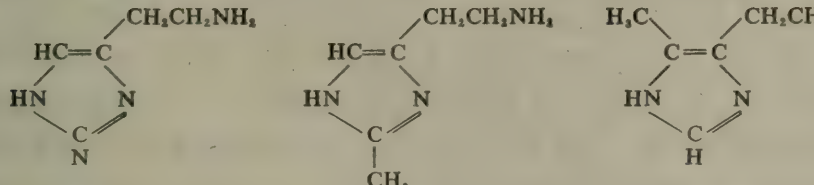


图 8-1 组胺及其衍生物分子图

实验表明[8]4-甲基组胺对 H_2受体的活性为组胺的 40%，对 H_1受体仅为组胺的 1/500。而 2-甲基组胺又显示了与 4-甲基组胺的活性差异，其对 H_2受体的作用仅为组胺的 2%，而对 H_1受体作用却为组胺的 17%。(见表 8-1)

另一个说明构象重要性的明显例子是肾上腺素类药物[9]。如图 8-2 中，化合物(1)和(3)，它们具有不同的药理作用，化合物(1)是 β-阻断药，而(3)却兼有 α-肾上腺素的兴奋作用和 β-阻

表 8-1 组胺及衍生物相对活性比较

化合物	相对活性	
	H_1 受体	H_2 受体
组胺	100	100
2-甲基组胺	17	2
4-甲基组胺	0.2	40

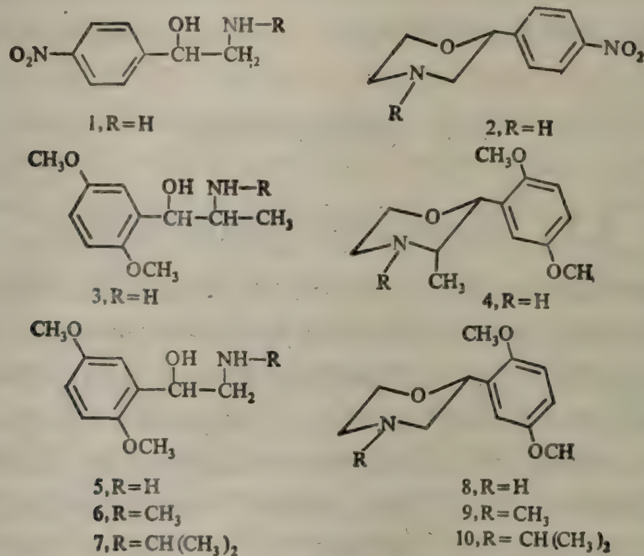


图 8-2 肾上腺素类药物及其吗啉类似物

断作用。当将吗啉环加入其乙基胺部分，合成如(2)和(4)化合物后，则原有的空间构象（指侧链与苯环的空间构象）发生了较大的改变，遂成为产生不同药理作用的主要原因之一。实验表明[10]，吗啉类似物(4)非但没有任何兴奋作用，相反地却显示出一定的α-受体阻断活性。

常用药物青霉素或头孢霉素的抗菌药理作用是众所周知的。它们之所以产生抗菌作用，是由于抑制细菌制造细胞壁所致。细胞壁的构成依赖于一些多糖链和多肽链的互相交织。在这种生化过程中，一条多个D-构型丙氨酸组成的肽链将加到另一条由几个甘氨酸组成的肽链。这种生化反应在转肽酶和羧肽酶的催化下才得以完成。丙氨酸肽链首先与这两种酶上的相应基团(即受体部位)相组合，然后再经催化缩合，使细菌得以继续生长。经过研究发现青霉素，包括半合成青霉素的构象(见图 8-3 I_a)恰好与D-构型丙氨酰D-构型丙氨酸(II)的构象(II_a)相吻合[11,12]。因而当青霉素进入体内后，它也可在酶的同样部位(受体)上结合，于是就不容置疑地占去了D-构型丙氨酸肽链的结合部位，这就阻止了其与甘氨酸链的缩合，使保护细胞的细胞壁不能完整构成，危及细菌的生长，从而达到抗菌的药理目的。

头孢菌素类(III)、硫霉素(thienamycin，IV)以及结构更为简单的化合物(V)也有很强的抑

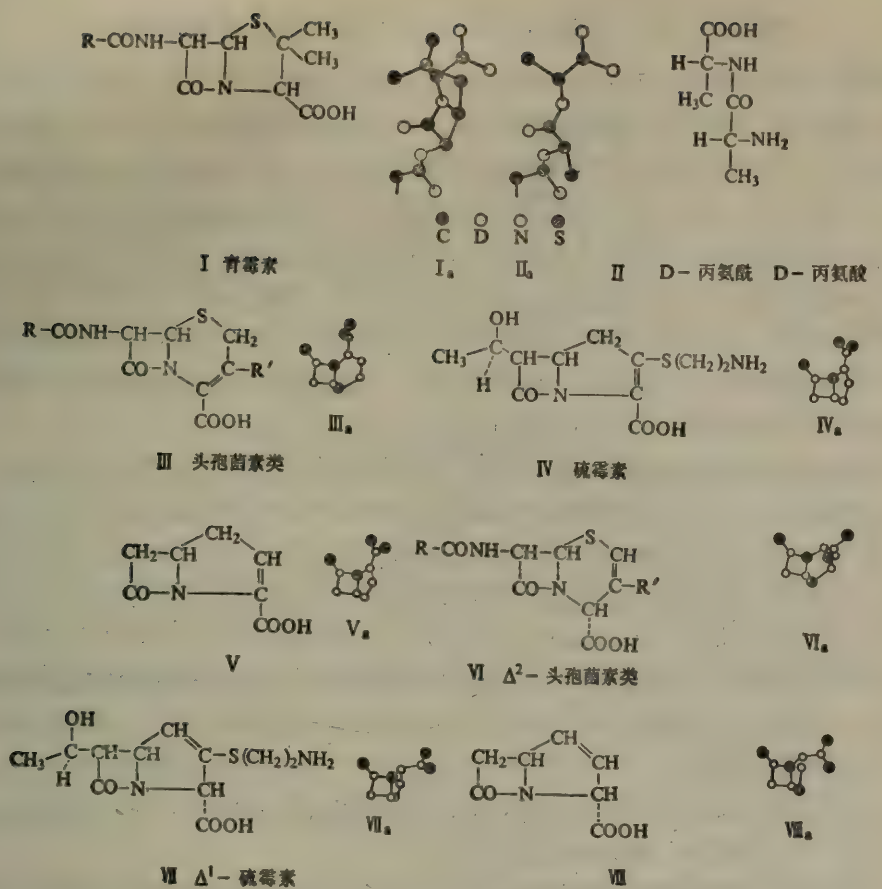


图 8-3 青霉素等分子结构及相应构象示意图

菌作用。但是，当将它们的化学结构略加修饰，如将头孢菌素环内双键位置由第 1 位移至第 2 位，硫霉素环内双键位置移至第 1 位，同样化合物 V 环内双键亦移至第 1 位，分别合成了图 8-3 中的 VI、VII、VIII 三个化合物，经试验表明，它们已丧失了抗菌作用。研究结果[13]指出，造成药理作用迥然不同的原因就是由于双键位置的迁移，同时导致环上羧基的构象发生了变化。由I_a、III_a、IV_a、V_a 的构象图象可以清楚地看到，有抗菌活性的青霉素、头孢霉素、硫霉素等的羧基占有环上的假平伏键（见 I_a 右上角平伏的 COO 三个原子和 III_a 等右上角三个原子）。反之，双键位移后的化合物 Δ^2-头孢菌素类（VI_a），Δ'-硫霉素（VII_a）及 $VIII_a$ 的羧基都处在环上假直立键。经研究，假平伏键的羧基与内酰胺基靠近，酰胺的氧原子与羧基的碳原子间相距约为 0.30～0.39 nm；相反，假直立键由于这两种基团相距较远，致使酰胺基上的氧原子与羧基碳原子间距在 0.41 nm 以上。正由于这个构象差别，造成假直立键化合物不能结合在前述酶上同一关键部位，不能抑制细菌的转肽酶或羧肽酶的催化缩合，从而失去了抗菌活性。

从以上几个例子中充分证明药物分子组成的空间构象支配着药物分子与受体相互作用的部位，可见其对产生药理作用的重要性。

8.2.2 分子轨道法在研究药物分子构象中的应用

对药物分子构象的研究表明，分子轨道法通过不同近似求解 Hartree-Fock-Rothaan 方程，得到药物分子的能量，从而可以推测药物的构象。这是因为在不同构象中，各个相互作用的原子或基团间距离不一，从而导致整个分子能量的差异。换言之，当分子内某一化学键旋转时，由于原子核与电子间的引力，以及各原子核之间和各电子之间的斥力，同时也由于价键和非键作用力的改变，使分子内旋转不同角度造成化合物具有不同的能量。能量较低的构象比能量较高的构象更为稳定，高能量的构象可能自动转化为低能量的构象。同样，与药物作用的生物大分子，也稳定在低能量的构象。而一系列类似药理作用的化合物当应与同一受体相作用，与受体结合的药效基团也应有相似排列，这已被许多研究所证实。因此分析若干类似作用化合物的较低能量构象，可以推测其可能的药效构象，即“活性构象”。

目前计算药物分子构象的量子生物学方法，有量子化学分子轨道法和分子力学方法。前者主要是从头计算法（ab initio），半经验分子轨道法的 MNDO、MINDO/3、INDO、CNDO/2，以及 PCILO 法和 EHMO 法，近来研究构象的分子轨道方法主要是从头计算法、MNDO 法以及 PCILO 法。分子力学方法是近年来应用于药物结构活性关系研究的常用方法之一，计算药物分子构象的分子力学方法主要是 Allinger 开创的 MM 系列中的 MM 2 和 MMP 2 程序[14]，还有 Adrewes 建立的 COMOL 计算系统[15]。为了便于计算药物分子与核酸结合的共同构象Pullman 创立了用从头计算碎片加和的分子力学方法——SIBFA 和 SIBFA 2[16,17]。它们已被越来越广泛地应用于量子药理学研究。

分子轨道法计算完全可以象处理几何形状那样处理构象。对于分子的一部分相对另一部分的一系列位置，可以计算并比较各种位置的能量。

若分子中只能绕一个单键发生旋转，或可以简化为一个单键的旋转，则可以用能量对角度变化的曲线形式描绘出构象能量变化的结果。

吴吉安等[18]在抗心律失常新药常咯啉及其类似物的量子药理学研究中，用 EHMO 法对如图 8-4 的常咯啉（4-{3′,5′-双[（N-吡咯烷基）甲基]-4′羟苯胺基}喹唑啉）分子进行了能量优选，以期找出能量最低的分子构象。

在进行能量优选计算时，注意到常咯啉分子 τ_1、τ_2 旋转与 τ_1' 和 τ_2' 旋转的对称性，以及 τ_3 和 τ_4 旋转的空间障碍，故采用近似处理的优选办法。在寻找出 $\tau_1\tau_1'$ 和 $\tau_2\tau_2'$ 对应的二面角 $\phi_1=\phi_1'=120°$，$\phi_2=\phi_2'=180°$ 为该局部能量最低的前提下，分别单独旋转 τ_4 和 τ_3。τ_4 对应的两面角 ϕ_4（N_{46}，C_{45}，N_1，C_2）从 120°始进行一次 EHMO 计算，而后每旋转 30°作一次计算，所获得分子轨道总能量 $E_{总}$ 列于表 8-2 中，依据表 8-2 给出的能量值，即可画出能量对角度的曲线图（图 8-5），由图可见 $\phi_4=0°$能量最低。

最终获得常咯啉能量最低的构象为 $\phi_1=120°$，$\phi_1'=120°$，$\phi_2=180°$，$\phi_2'=180°$，$\phi_3=90°$，$\phi_4=0$。

在药理学研究中，绕单键旋转的能量优选毕竟不多，上述的例子也仅是作为较大分子的过渡计算。对药物分子而言，往往有几个单键同时旋转。在这种情况下，分子体系的能量应是所有

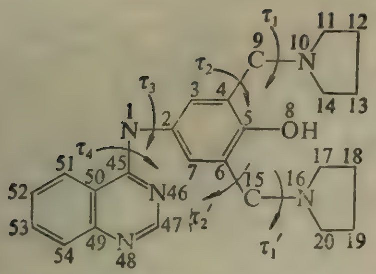


图 8-4 常咯啉分子 Dreiding 模型旋转图

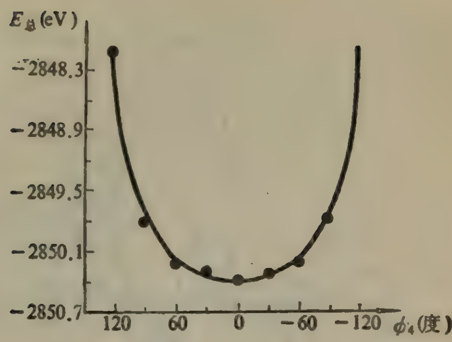


图 8-5 τ_4 旋转能量曲线

表 8-2 τ_4 旋转能量计算值(eV)

ϕ_4 两面角	120°	90°	60°	30°	0°
$E_{总}$	−2848.1364	−2849.8056	−2850.1953	−2850.3266	−2850.3325

对应的可旋转单键的两面角 τ 的函数。药物分子的稳定构象即对应于该分子能量函数的最低值。

例如,在抑制性介质 γ-氨基丁酸(GABA)分子的构象中[19],一共可以进行 τ_0、τ_1、τ_2、τ_3 四个二面角的单键旋转(图 8-6)。GABA 的总能量可看作是这些二面角的函数。亦即

$$E_{总}=E(\tau_0、\tau_1、\tau_2、\tau_3)$$

按照通常量子化学计算中确定分子空间坐标所沿用的二面角定义(参见第 2 章),可方便地从图 8-6 中找到 τ_i 所对应的原子。如 τ_1 即为原子 C_1,C_2,C_3 和 C_4,C_2,C_3 分别构成的平面的夹角,它们的方向规定为右手法则。亦即从 C_2 往 C_3 看,以 C_2C_3 为转轴,顺时针旋转时二面角为正值,反时针为负值。$\tau_1=0$,对应于 C_1C_2 与 C_3C_4 键同平面且顺式排列,若同平面反式排列,则相应 $\tau=180°$。于是 GABA 分子中四个两面角分别为 $\tau_0(0,C_1,C_2,C_3)$,$\tau_1(C_1,C_2,C_3,C_4)$,$\tau_2(C_2,C_3,C_4,N^+)$,$\tau_3(C_3,C_4,N^+,H)$。

由于 4 个两面角同时旋转,给计算和结果分析带来了很大困难。为了减轻这些困难,往往选择影响构象最主要的两个二面角进行旋转计算,而对其中比较次要的二面角则按有机化学的经验取固定值。在 GABA 中,把羧基的两个氧原子取与 C_2C_3 共平面,即 $\tau_0=0°$,τ_3 按正四面体的标准构型取与 C_3C_4 的交叉位构象,即 $\tau_3=\pm60°$。于是把 GABA 分子能量化简为 τ_1、τ_2的函数。计算分子 τ_1、τ_2 从 0°～360°或者从−180°～180°变化的能量值(角度变化的步长视不同对象而定)。计算结束后,求出能量E的最小值E_{min},并计算旋转 τ_1、τ_2 不同角度所得能量与 E_{min} 的差值,便可绘制用等高线图形所描绘的能量、二面角 τ_1 和 τ_2 的二维图形。在 GABA 分子计算中,以 τ_1 作纵坐标,τ_2 为横坐标,对应于某一对 $\tau_1\tau_2$ 角度的相对能量值在图中以等高线画出。结果如图 8-7,从图 8-7 可以看出,E_{min}相对值为零的区域对应于$\tau_1=\pm60°$,$\tau_2=0°$,表明这是GABA

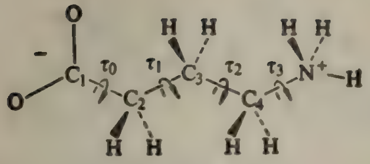


图 8-6 GABA 分子旋转示意图 $\tau_0=\tau(0,C_1,C_2,C_3)$ $\tau_1=\tau(C_1,C_2,C_3,C_4)$ $\tau_2=\tau(C_2,C_3,C_4,N^+)$ $\tau_3=\tau(C_3,C_4,N^+,H)$

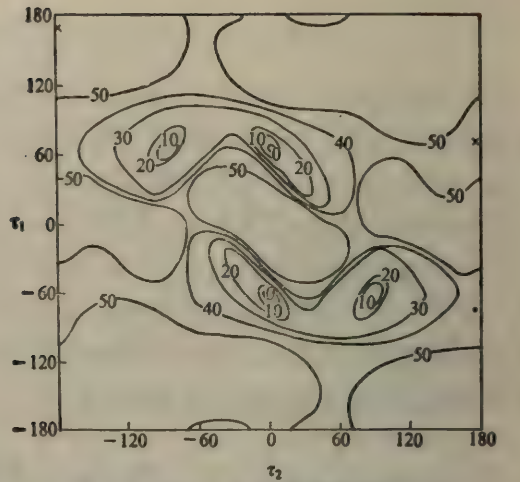


图 8-7 GABA 分子的构象图(PCILO 法)

分子的两个稳定构象区。而E_{min}相对值≈10 的区域,对应于$\tau_1=\pm60°$,$\tau_2=\pm90°$,显示GABA分子还存在两个次稳定构象区。

构象图的获得给药理学家了解药物分子的构象带来了很大的方便，并提供了稳定构象的确切信息。严格地讲，精确的构象图应通过从头计算法获得。但是在目前，由于受计算条件的限制,对一些较大的药物分子仍使用一些半经验的分子轨道法或分子力学方法获得构象图。不过,需要指出的是,运用从头计算法获得构象图的趋势正在与日俱增。

单个构象图的分析有时会带来一定的偏差。这主要是由于它忽略了热力学的论点，即对药理效应的影响。药物分子的活性构象不一定为具稳定构象或称平衡构象[2]。药物分子在平衡构象附近的非平衡构象 (non-Equilibrium conformation)往往起着重要药理作用[20]。这是因为热力学中的熵效应常常是重要的因素,考察势能面上两个势阱的相对分布,则可发现，它既和分子脱离势阱的难度如何有关,也须和势阱宽度不同造成分子进入势阱的难易程度有关[1]。

图 8-8 中显示了一维和二维的构象图,它清晰地表明这种影响:标记为A的构象显示其势能面宽而势阱浅,而B恰恰相反,势能面上窄而势阱深。则构象A的分子分布会比内能占优势但所占几率低的构象B更多。

Richards 等用 ab-initio 法和 EHMO 法对组胺及 4-甲基组胺进行了构象研究,提出了组胺 H_1 受体的"必要构象"("H_1-essential" conformation),从而成为组胺量子药理学研究中较有影响的成果。

他们在用 ab initio 法采用 STO-3 G 基组对组胺及 4-甲基组胺进行计算后发现,组胺阳离子和 4-甲基组胺阳离子的内旋转势能的变化幅度,与用 EHMO 法获得的结果相吻合[21]。表 8-3 列出了两种方法计算的结果。可见两种计算方法的差值相同,均为 1.25 eV。可以认为,用较简便的 EHMO 法在研究组胺构象变化时，可代替复杂的 ab initio 法。通过 EHMO 法计算获得的构象能量图显示,组胺及 4-甲基组胺的构象图可分析两个分子变化范围的差别[22]。

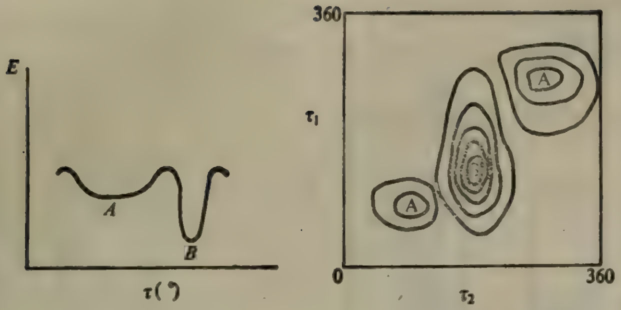


图 8-8 两种势能阱的比较

表 8-3 阳离子内旋转势能计算值（eV）

分子＼方法	EHMO	ab initio
组胺	0.14	0.24
4-甲基组胺	1.39	1.59

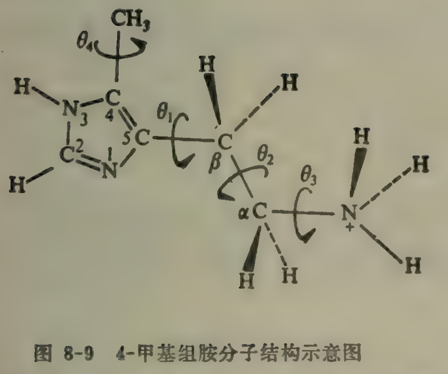


图 8-9 4-甲基组胺分子结构示意图

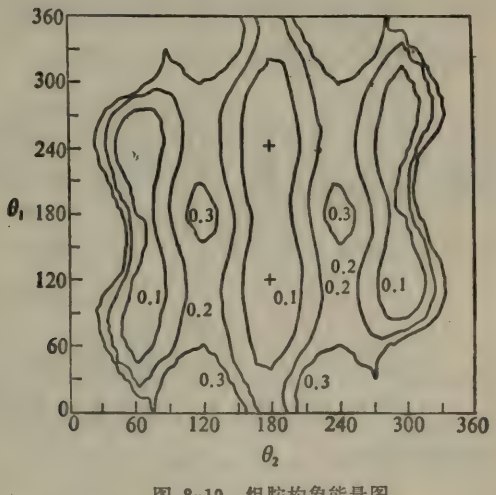


图 8-10 组胺构象能量图

4-甲基组胺有 4 个可以自由旋转的扭角，如图 8-9，按前述原则，在四个扭角中 θ_4 和 θ_3 按标准构型，取能量最低的扭角为 $\theta_4=120°$，θ_3 取交叉位（$\theta_3=60°$），这样分子能量简化为 θ_1 和 θ_2 的函数。计算的组胺和 4-甲基组胺构象能量图为图 8-10、图 8-11。

分析比较图 8-10 和图 8-11 发现，两者同样高于总体能量最小值 $E_{min}=0.1$ eV数值的范围，组胺的 θ_1 和 θ_2 变化范围大于 4-甲基组胺。若将图 8-10 与图 8-11 重叠，仅仅取 0.1 eV 的等高线作图，见图 8-12。

图 8-12 表明，4-甲基组胺 0.1 eV 等高线在虚线以内，而组胺则在实线以内。阴影部分表明

该区域 4-甲基组胺的 θ_1、θ_2 不能达到，相反组胺却能达到。由上述分析可以认为，组胺分子的 H_1 受体的活性是 4-甲基组胺的 500 倍，是由于对 H_1 受体要求组胺分子的侧链 θ_1 和 θ_2 取如图 8-12 中阴影部位的构象。这部分构象称之为H_1受体的“必要构象”。

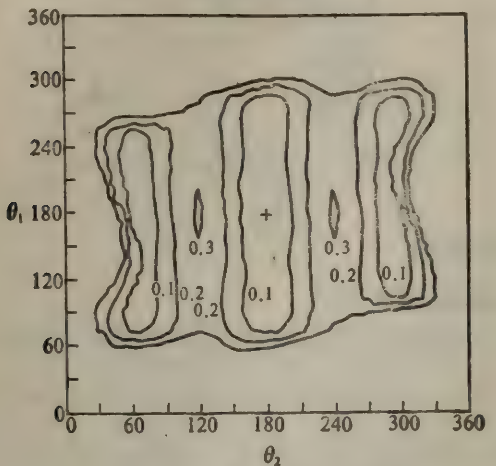


图 8-11 4-甲基组胺构象能量图

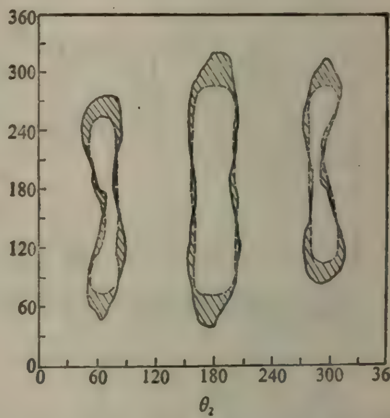


图 8-12 组胺和 4-甲基组胺构象图的重叠

但是H_1受体的必要构象却不是组胺的最稳定构象，而是偏离稳定构象的某个较不稳定构象。这就充分说明当药物小分子与受体大分子作用时，产生生物活性的“活性构象”有时是非平衡构象。

由于分子所取的构象具有统计性质，因而简单地用构象图是难以表达能量的区域变化的。加上把热力学中熵的改变一同考虑，即可把能量图化成 Boltzman几率分配函数。这个函数可定义为

$$Z_{构象}=\sum_{\tau_1\tau_2} e^{-\frac{E_{\tau_1,\tau_2}}{KT}} \tag{8.1}$$

其中E_{τ_1,τ_2}是取不同 τ_1、τ_2 角度所算得的能量，K是 Boltzman常数，T为绝对温度(通常取 37°C)。求和是对一系列点进行的。把几率函数对整个构象表面进行积分可得到归一化因子 Z，并运用下式校正所有的点

$$Z^{新}_{\tau_1,\tau_2}=Z^{旧}_{\tau_1,\tau_2}\Big/Z \tag{8.2}$$

从而使函数积分为 1，再对 $Z^{新}_{\tau_1,\tau_2}$ 作图，即可获得几率分布图。若用 Z 对指定区域积分可得该区域几率，将几率乘以分子数可得该区域的集居数（population)。同时可以画出分布图象。若对所有神经介质都采用两个扭角旋转的情况，便能画出集居数图，见图 8-13。

从这类图象可以看出，分子具有某几率处于该构象的区域，使它有更多的机会与受体作用。而对同系列分子进行比较，则可画出每个化合物的 99%的界线图，这样，分析其构象作用就方便得多了。

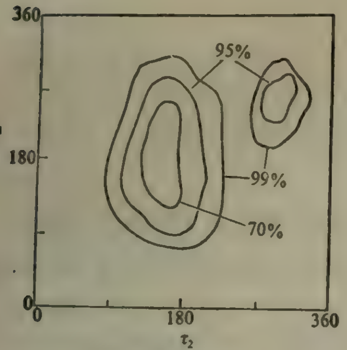


图 8-13 集居数图

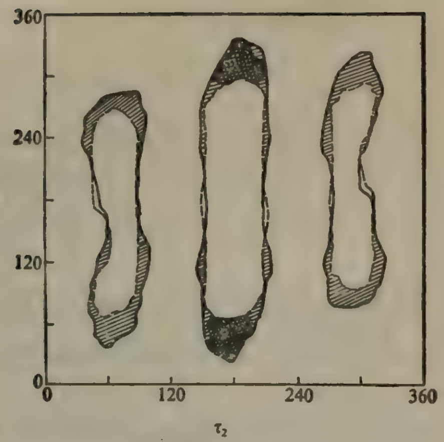


图 8-14 构象百分率图象的对比

图 8-14 是构象百分率图象对比，画点的面积表示适用于组胺不适用于 4-甲基组胺 的区域。而黑色面积在所有 H_1 活性激动剂的 99%界线内，代表这类活性的“必要构象”。

由于分子的非平衡构象不可能通过 X 衍射或 NMR 测得，而分子轨道计算对平衡构象和非平衡构象同样有效，因而量子药理学在构象研究中显示了 其 独特的重要作用，并将愈来愈引起人们的注意。

8.2.3 去甲肾上腺素的构象研究

去甲肾上腺素及乙酰胆碱分别是两类不同的神经系统的神经冲动 传 递 媒 介，称为神经递质。由于分泌的神经递质的不同，近年来将这两类神经称为肾上腺素能神经 与 胆碱能神经。肾上腺素能神经是指释放去甲肾上腺素为媒介的神经纤维。由于去甲肾上腺素在药理学中的重要地位，因而用分子轨道法对其进行研究也引起了广泛的兴趣。

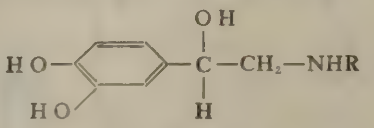
在构象研究中，不仅药物有其“活性构象”，同样与药物相互作用的生物体大分子(如受体)，也需稳定在低能量的构象的空间位置。由于生物大分子化合物结构复杂，有时可能有几种不同构象的能量都较低。在受到外界影响时，由于互变的能量障碍不大，则几 种 构象就可能同时存在。当药物小分子或内源活性物质与这些大分子的不同构象相作用时，将分 别 产生不同的药理效应，这称之为一种受体的几种受体亚型(有关受体亚型在下节还将叙述)。

去甲肾上腺素可与神经突触后受体结合，产生效应，该受体有几种受体亚型。

早期的量子药理学计算表明，去甲肾上腺素分子侧链上两个碳原子旋转在不同角度时，以氨基与苯基成对位交叉式的能量最低，氧原子与氮原子相距为 0.286 nm。麻黄碱的计算也有相似的有利构象，氮原子与氧原子距离为 0.295 nm。而假性麻黄碱为邻位 交 叉式，意味着与同一受体的结合不利，故而拟交感作用很弱。

进一步研究表明，去甲肾上腺素的受体有 α 、β 亚型，例如与 α 受体结合产生收缩血管、升高血压等效应。与 β 受体结合产生增快心搏、增强心肌收缩并舒张支气管等效应。

把去甲肾上腺素的化学结构局部加以变动(改变R取代基),则对这二种受体亚型结合的



Xa. R═H去甲肾上腺素

b. R═CH_3肾上腺素

c. R═$CH(CH_3)_2$异丙肾上腺素

d. R═$C(CH_3)_3$

弱随即发生变化,例如去甲肾上腺素(*Xa*)以作用于α受体为主,而氨基上带有甲基的肾上腺素(*Xb*)对α和β受体都有作用,若氨基上取代有异丙基(*Xc*)或叔丁基(*Xd*)时,则主要作用于β受体。

为研究这一不同受体亚型的作用,探索庞大取代基怎样影响其对受体的亲和力,有人用称之为ISC-PEM-MO的属于分子力学原理的方法和程序对*Xa*和*Xb*分子进行分子构象的优化计算[23]。将如图8-15所示的两个分子键长键角按标准数据固定,旋转相应的二面角,画出两两对应的能量构象图。按其能量分布的情况可形象地归结为如图8-16所示的三种构象。

在图8-16中,*A*式能量为最低,*C*略高,比*A*高出2.5 kcal/mol,*B*能量最高,比*A*大3.5 kcal/mol。对这些构象计算结果的分析表明,*A*中苯基与铵基相距最远,系对位交叉式,位阻最低,去甲肾上腺素的分子若取*A*构象,则羟基与胺基间可形成分子内氢键,更能降低能量,因而*Xa*分子取对位交叉构象,也是*A*受体的必要构象。而*C*构象的铵基在苯基与羟基之间,*B*构象的去甲肾上腺素不可能形成分子内氢键,*C*构象处于氨基、羟基、苯基都是邻位交叉的部位,能量升高,经分析肾上腺素分子的构象能量图,发现处于*A*构象的面积比去甲肾上腺素的面积减少,这说明氨基上引进取代基会减低对位交叉构象*A*的存在,指示氨基引进庞大取代基团,会使构象自*A*趋向*C*,构象发生变化,与之适应的受体亚型就不同,药理效应随之发生变化。

作者还使用获得的优势构象,对*Xa*和*Xb*分子进行量子化学CNDO/2计算,发现氨基上

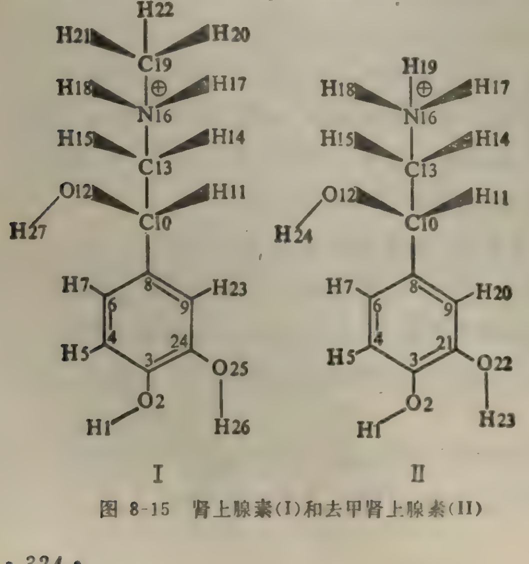


图 8-15 肾上腺素(I)和去甲肾上腺素(II)

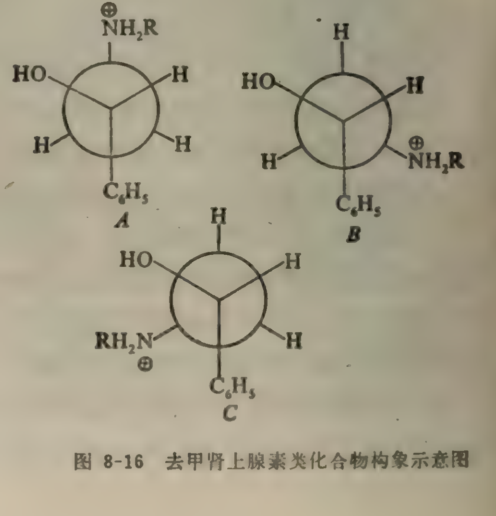


图 8-16 去甲肾上腺素类化合物构象示意图

若引进取代基，将使氮原子上电荷密度降低($Xa:QN=4.9661$；$Xb:QN=5.0107$)，表明氮原子的电荷密度大小亦可能与受体不同亚型有关。

8.2.4 多巴胺构象的研究

多巴胺与去甲肾上腺素的化学结构十分相象，只是在侧链上少了一个羟基，但计算结果以邻位交叉式能量为最低，这是由于侧链上没有羟基，就不存在去甲肾上腺素分子侧链上羟基上氧原子的孤对电子与铵离子间产生的引力。多巴胺与去甲肾上腺素构象既然不同，与之相作用的受体也将不同。

多巴胺的研究在药理学中也是相当重要的，我们知道，无数震颤麻痹症药物和精神病药物都作用于多巴胺能神经受体。如抗精神病药物在多巴胺受体上主要起拮抗剂的作用，代表性的药物有氯丙嗪、氟哌丁苯、吗啉吲酮(molindone)、哌咪清(pimozide)、Z-三氟噻吨(Z-flupenthixol)、克噻平(ioxapine)、丁克吗(butaclamol)、异丁克吗和异丙克吗(dexclamol)等。

经X衍射、NMR测定提示，由于其具有揉曲侧链，所以多巴胺的构象是易变的。经量子化学INDO法计算优势构象[24]，对于顺式和反式共有十种类似最低能量的构象，从而也证实多巴胺构象的易变性。

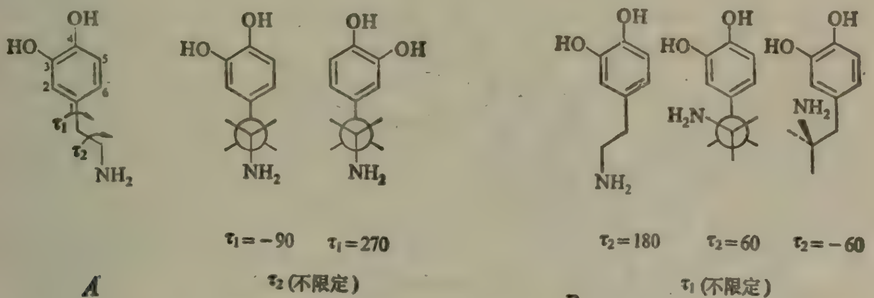
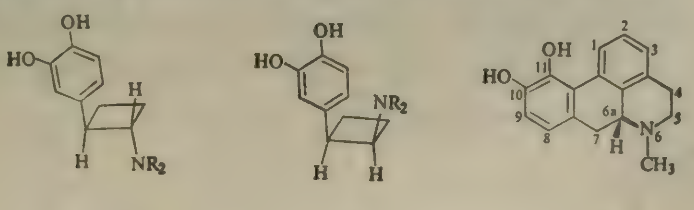


图 8-17 多巴胺构象。(*A*) 为多巴胺分子构象旋转示意。(*B*) 为可能的十种构象。

图 8-17 所示，*a* 是多巴胺分子构象τ_1、τ_2旋转示意图。*b* 表示二类扭角范围很宽的构象，1) τ_2不限定，$\tau_1=-90°, 270°$，2) τ_1不限定，$\tau_2=180°, \pm 60°$。因为它们之间能量差别很小，故其中任何一种构象在相应的受体上都可呈现生物活性。

采用刚性的环丁烷试验，反式化合物(I)代表延伸型多巴胺构象，顺式化合物(II)代表折叠型的多巴胺构象。计算时τ_1不受限制，τ_2为0°或120°开始，结果表明伸展型($\tau_2=180°$)的形式是多巴胺活性最大的构象，与受体结合试验也表明反式化合物的活性更高[25]。



I　　II　　III

有多巴胺能神经激动剂和拮抗剂。多巴胺激动剂研究得最多的也许是阿扑吗啡(III)，它几乎是一平面状的四环叔胺化合物，分子中含有扭角 $\tau_1=40^\circ$，$\tau_2=180^\circ$ 多巴胺的结构部分。而丁克吗((+)-butaclamol)(IV)，它们均属刚性化合物，经 X 衍射并辅以 INDO 计算，表明它们具有延伸型多巴胺构象，图 8-18 显示了三个化合物都带有相同的药效基团。

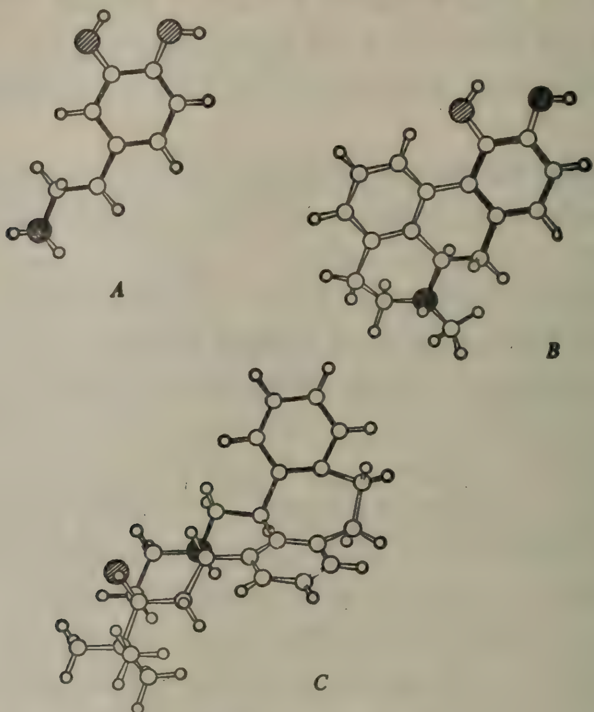
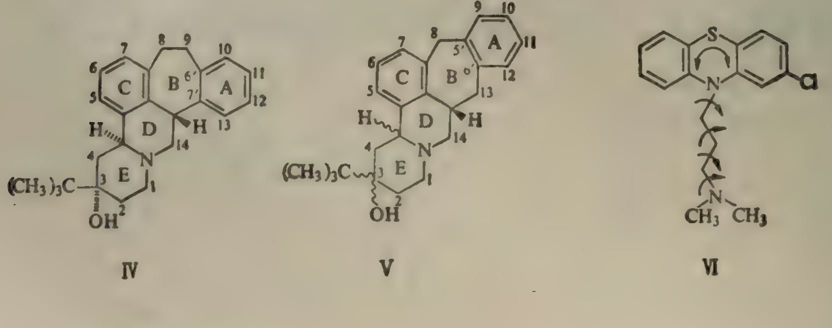


图 8-18　多巴胺(*A*)阿扑吗啡(*B*)丁克吗(*C*)构象

虽然量子药理学计算表明，当多巴胺的 CH_2NH_2 侧链与芳环处于共平面或相互垂直这两种构象时，其势能相差不多，因而都可能存在，但刚性结构的多巴胺激动剂如阿扑吗啡，或 2-氨基四氢萘等都是近乎平面的化合物，这表明它们以平面构象亦即多巴胺延伸型构象与受体结合。在 2-氨基四氢萘分子内，环己烷为半椅式，氨基处于平伏键，这种构象与阿扑吗啡相近。多巴胺



Ⅳ　　Ⅴ　　Ⅵ

侧链旋转,可得 α 旋转体和 β 旋转体，这两者羟基与氨基氮原子与氧原子距离不同。合成带刚性结构的 5,6-二羟基与 6,7-二羟基-2-氨基四氢萘代表 α、β 旋转体构象，经比较一系列化合物抑制大鼠自发运动活力，α 旋转体多巴胺延伸型构象的活性要强得多[26]。

在多巴胺能拮抗剂丁克吗(布他拉莫)类似物的合成及药理研究中[27]，发现凡不带多巴胺活性构象的均无活性。但异丁克吗((+)-isobutaclamol)分子(V)内并不带有苯乙胺成分，却有苯丙胺结构，其活性与丁克吗相近。

在化合物(IV)和(V)中，E 环为椅式构型，叔丁基处平伏键位，B 环稍有卷曲。按照 X 衍射测得的构象，两者不能重合，但取量子药理学计算中所得能量较低的一种构象，(A 构象与 B 构象，取 B 构象)，可见两分子中药效基团重叠，并与阿扑吗啡的构象一致，即具有延伸型多巴胺构象。见图 8-19、图 8-20。

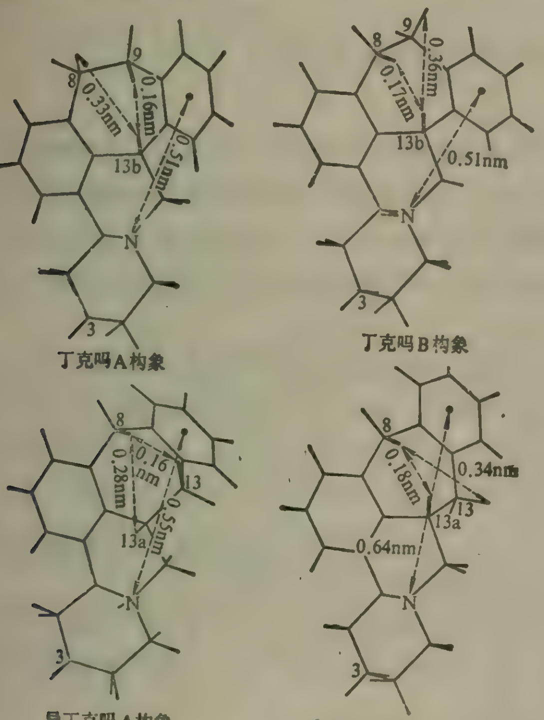


图 8-19 化合物 IV和 V 的 A 构象与 B 构象

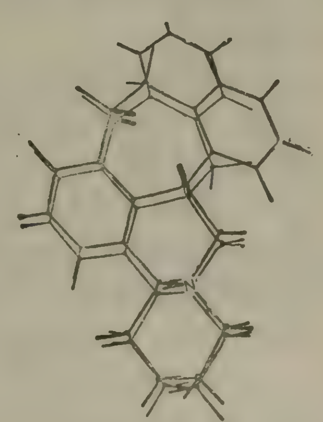


图 8-20 化合物 IV、V的 B 构象重叠图

这种构象分析还用于其他结构看起来显然不同的若干精神病药物的研究上。图 8-21 展示了丁克吗的类似物异丙克吗(dexclamol)与氯丙嗪(VI)的构象。这是 Tollenaere 等应用 X 衍射和分子轨道 PCILO 法计算进行研究的结果[28]，表明了两者药效构象的一致性。

8.2.5 抗生素类抗肿瘤药物的量子药理学研究

自从发现放线菌素(cactinomycin)对何杰金氏病有疗效以后，人们开始从微生物中寻找抗肿瘤药物，业已成为抗肿瘤新药的主要来源之一。

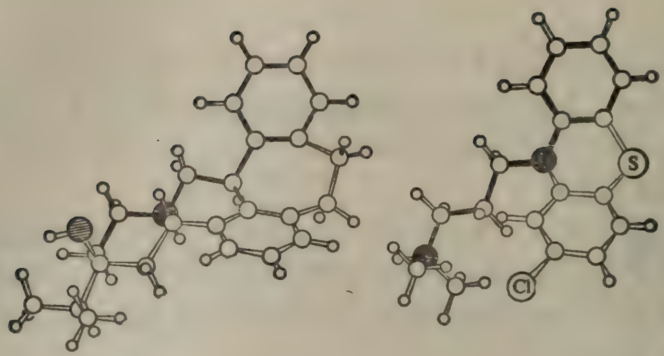


图 8-21 氯丙嗪(*B*)与(+)-declamol(*A*)分子药效构象图

由于分子生物学的发展，对这类抗癌药物的作用机制有了比较深入的认识。抗肿瘤抗生素的作用大部分是抑制 DNA、RNA 和蛋白质的合成。对于具有醌式的平面芳香结构的抗生素来说，如柔红霉素(daunorubicin)(图 8-22)、阿霉素(adriamycin)等，这类化合物对 DNA 分子的结合，是发生嵌插作用，阻断产生依赖于 DNA 的 RNA，即抑制转录过程，从而阻止了蛋白质合成而使细胞受损或死亡。

正因如此，药物分子的核酸嵌插作用引起了量子生物学研究的广泛兴趣。药物分子与核酸的嵌插相互作用是指分子结构中平面的芳香色团部分平行地嵌入到核酸相邻碱基对之间，其余的部分安放在双螺旋的沟槽内。研究嵌插相互作用分实验方法和理论方法，理论方法可分成两大类，一是以 X 光实验结果为起点，如 Neidle[29]、Kollman[30] 等人对 X 光衍射的结果，进行复合物构象的能量优化处理；Pullman 等[31-37]以晶体中一核酸嵌插受体为模型，对柔红霉素等进行 DNA 的亲和性和选择性研究；另一类是从理论核酸模型入手进行研究。

陈凯先等[31,32] 运用 Pullman 发展的量子化学与分子力学结合的分子内和分子间相互作用的计算方法——SIBFA 和 SIBFA 2[16,17](见第 2 章分子力学的有关介绍)，计算了阿霉素和柔红霉素的构象能，对它们的嵌插选择性作了理论计算。

相互作用能的计算：

分子内构象能

$$\Delta E_{conf.} = E_{MTP} + E_{pol} + E_{rep} + E_{disp} + E_{tor} \tag{8.3}$$

分子间相互作用能

$$\Delta E_{inter} = E_{MTP} + E_{pol} + E_{rep} + E_{disp} + E_{CT} \tag{8.4}$$

上两式中，E_{MTP} 为静电势作用能，E_{pol} 为极化作用能，E_{rep} 为排斥作用能，E_{disp} 为色散作用能，E_{tor} 为扭转作用能，E_{CT} 为电荷转移能。

柔红霉素与双螺旋低聚核苷酸(double-stranded polynucleotides)的复合物的生成可以有六种结合形式(图 8-23)

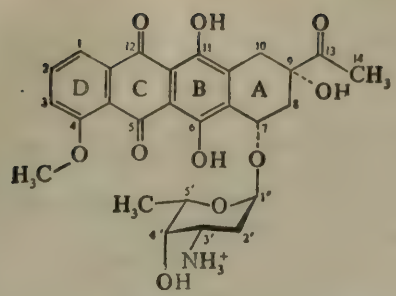


图 8-22 柔红霉素(daunomycin)分子图(DNM)

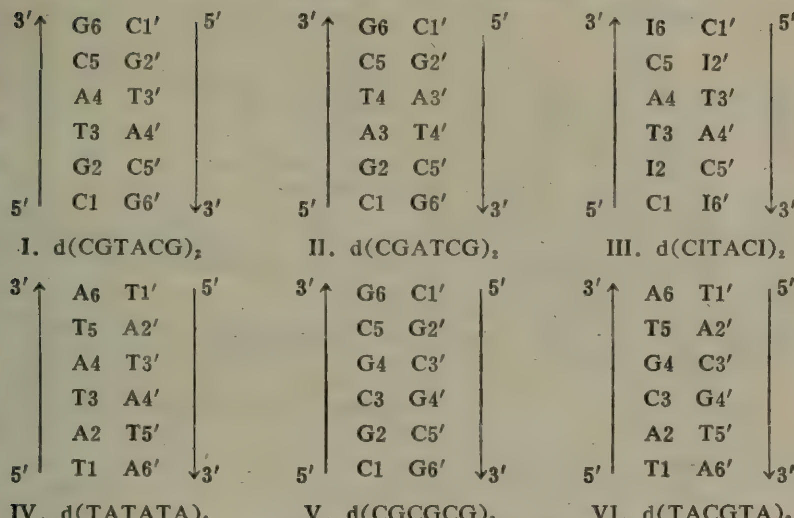


图 8-23 六种结合形式双螺旋低聚核苷酸碱基编码

经过计算 DNM-DNA 六种形式的复合物的能量数据列于表 8-4 中，稳定化能$\delta E=\Delta E_{\text{inter.}}+\Delta E_{\text{conf.}}+\Delta E_{\text{unstack}}$。从 δE 值的大小判断，其稳定顺序为

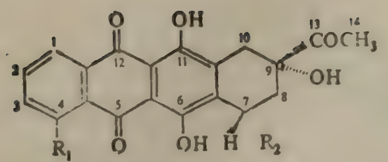
$$\underset{\text{II}}{d(CGATCG)_2}\geqslant\underset{\text{I}}{d(CGTACG)_2}>\underset{\text{IV}}{d(TATATA)_2}\geqslant$$

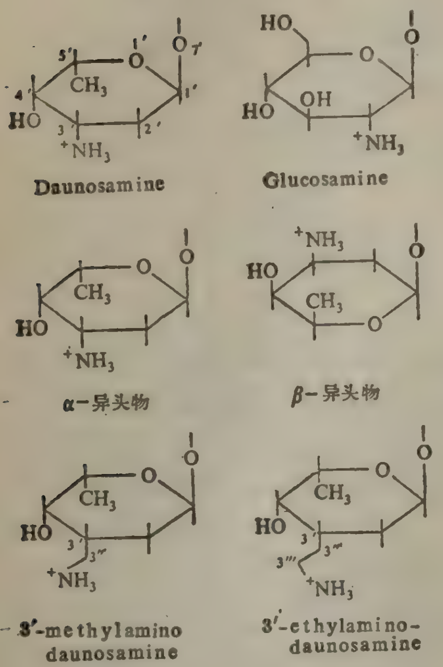
$$\underset{\text{III}}{d(CITACI)_2}\gg\underset{\text{V}}{d(CGCGCG)_2}>\underset{\text{VI}}{d(TACGTA)_2}$$

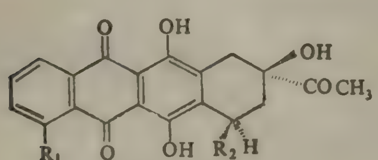
从上述顺序可知，DNM-DNA 的嵌插选择性是polyd$(TA)_2$的结合能比对 polyd$(CG)_2$强。d$(CGATCG)_2$与 d$(CGTACG)_2$相当。作者以 d$(CGATCG)_2$-DNM 复合的形式，进一步运用 SIBFA 计算方法，计算了 DNM 的 8 个衍生物与双螺旋低聚核苷酸稳定构象所对应的各种能量。

图 8-24 显示了他们所进行计算的八个分子的结构，值得指出的是，化合物 7 和 8 是作者设计的新衍生物，其结构特征是 DNM 分子中糖基上 3′ 位的$-NH_3^+$基团改变为$-CH_2-NH_3^+$(7)或$-CH_2CH_2NH_3^+$(8)。

所有衍生物与双螺旋低聚核苷酸碱基结合的编号见图 8-25。







7,9-双异构体

N°	Compounds	R_1	R_2
1	Daunomycin	$-OCH_3$	Daunosamine
2	4-demethoxydaunomycin	—H	Daunosamine
3	β-anomer of 4-demethoxy-daunomycin	—H	Daunosamine
4	4-demethoxy-7,9-bis-epi-daunomycin	—H	Daunosamine
5	β-anomer of 4-demethoxy-7,9-bis-epi-daunomycin	—H	Daunosamine
6	Glucosamine analog of daunomycin	$-OCH_3$	Glucosamine
7	Methylaminedaunomycin	$-OCH_3$	3′-methylamine daunosamine
8	Ethylaminedaunomycin	$-OCH_3$	3′-ethylamine daunosamine

图 8-24 柔红霉素类似物结构

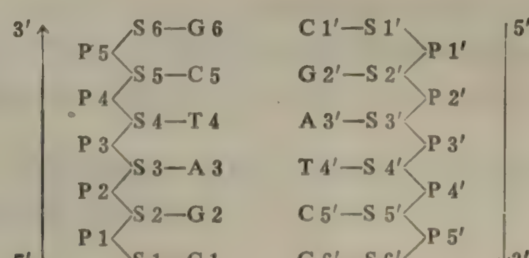


图 8-25 DNM 衍生物与双螺旋低聚核苷酸碱基结合编码

表 8-4 柔红霉素与 I—VI 复合物相互作用能(kcal/mol)

Hexanucleotide	d(CGTACG)$_2$	d(CGATCG)$_2$	d(CITACI)$_2$	d(TATATA)$_2$	d(CGCGCG)$_2$	d(TACGTA)$_2$
ΔE_{inter}	−433.5	−433.9	−429.7	−428.8	−416.4	−413.4
E_{MTP}	−369.2	−369.2	−368.5	−366.8	−353.6	−350.6
E_{pol}	−20.8	−20.3	−21.1	−21.5	−20.9	−21.1
E_{CT}	−14.0	−14.0	−10.8	−11.0	−15.2	−13.6
E_{disp}	−85.9	−86.4	−80.0	−82.8	−88.6	−86.3
E_{rep}	56.3	56.1	50.7	53.3	61.8	58.2
$\Delta E_{conf.}$	2.6	2.6	2.6	2.6	2.7	2.7
$\Delta E_{unstack}$	10.5	10.5	10.8	9.5	10.5	9.5
δE	−420.4	−420.8	−416.3	−416.8	−403.2	−401.2
δ	0.4	0.0	4.5	4.0	17.6	19.6
ΔE^*	−147.4	−147.6	−145.0	−143.6	−130.1	−127.1
E^*_{MTP}	−91.5	−91.4	−90.0	−89.0	−75.5	−72.6
δE^*	−134.3	−134.5	−131.6	−131.5	−116.9	−114.9
δ^*	0.2	0.0	2.9	3.0	17.6	19.6

表 8-5 列出 1—8 化合物与 d(CGTACG)$_2$ 优势构象所对应的各作相互作用能。

表 8-5 化合物 1—8 与 d(CGATCG)$_2$相互作用能 (kcal/mol)

	1 DM	2 4-Demethoxy DM	3 β-Anomer of 4-demethoxy DM	4 4-Demethoxy- 7,9-*bis*-epi DM	5 β-Anomer of emethoxy- 7,9-*bis*-epi DM	6 Glucosamine analog of DM	7 Methylamine DM	8 Ethylami- ne DM
ΔE_{inter}	−433.9	−435.5	−401.7	−383.7	−371.4	−411.4	−450.1	−453.1
E_{MTP}	−369.2	−371.8	−344.4	−350.1	−344.0	−350.6	−381.6	−380.7
E_{pol}	−20.3	−20.6	−18.6	−22.3	−19.3	−19.2	−21.9	−22.6
E_{CT}	−14.0	−13.4	−14.6	−9.8	−5.5	−14.6	−13.4	−15.2
E_{disp}	−86.4	−84.2	−75.6	−37.2	−46.0	−82.4	−87.1	−93.1
E_{rep}	56.1	54.6	51.5	35.8	43.2	55.4	53.8	58.4
ΔE_{conf}	9.7	9.8	3.4	8.4	8.1	8.0	7.5	23.6
δE	−424.2	−425.7	−398.3	−375.3	−363.3	−403.4	−442.6	−429.5
δ	0.0	−1.5	25.9	48.9	60.9	20.8	−18.4	−5.3
ΔE^*_{inter}	−147.6	−149.9	−118.9			−124.2	−160.7	−163.5
E^*_{MTP}	−91.4	−94.1	−71.3			−73.9	−100.9	−100.4
δE^*	−137.9	−140.1	−115.5			−116.2	−153.2	−139.9
$\overset{*}{\delta}$	0.0	−2.2	22.4			21.7	−15.3	−2.0
IC_{50}(μmol/l)	18[b]	18	38	50	>100			
K_a(l/mol)	3.3×10^{6}[c]	4.4×10^{6}						
	3.3×10^{6}[d]					$71.\times10^{4}$[d]		

从上表可知以 $\delta E=\Delta E_{inter}+\Delta E_{conf}$ 为稳定能,以它来衡量,其复合物的结合强弱为:

$$7>8>2\geqslant1(\mathrm{DNM})>6>3>4>5$$

表上的结合能顺序令人满意地与它们各自的抗肿瘤活性大 小平行 (见 IC_{50} 和 K_a)。令人振奋的是,预先设计的 7 和 8 两个新化合物,应具有比 DNM 强得多的抗肿瘤活性。这一理论研究表明,量子药理学对药物分子构象计算,至少在 DNA 嵌插模型作用机制的研究中，已进入定

量设计药物的水平。

图 8-26 显示了 DNM 和 7、8 分子与 $d(CGATCG)_2$ 复合稳定构象的空间分布图。这些图形采用 FIGATOM 作图程序[88]完成的。

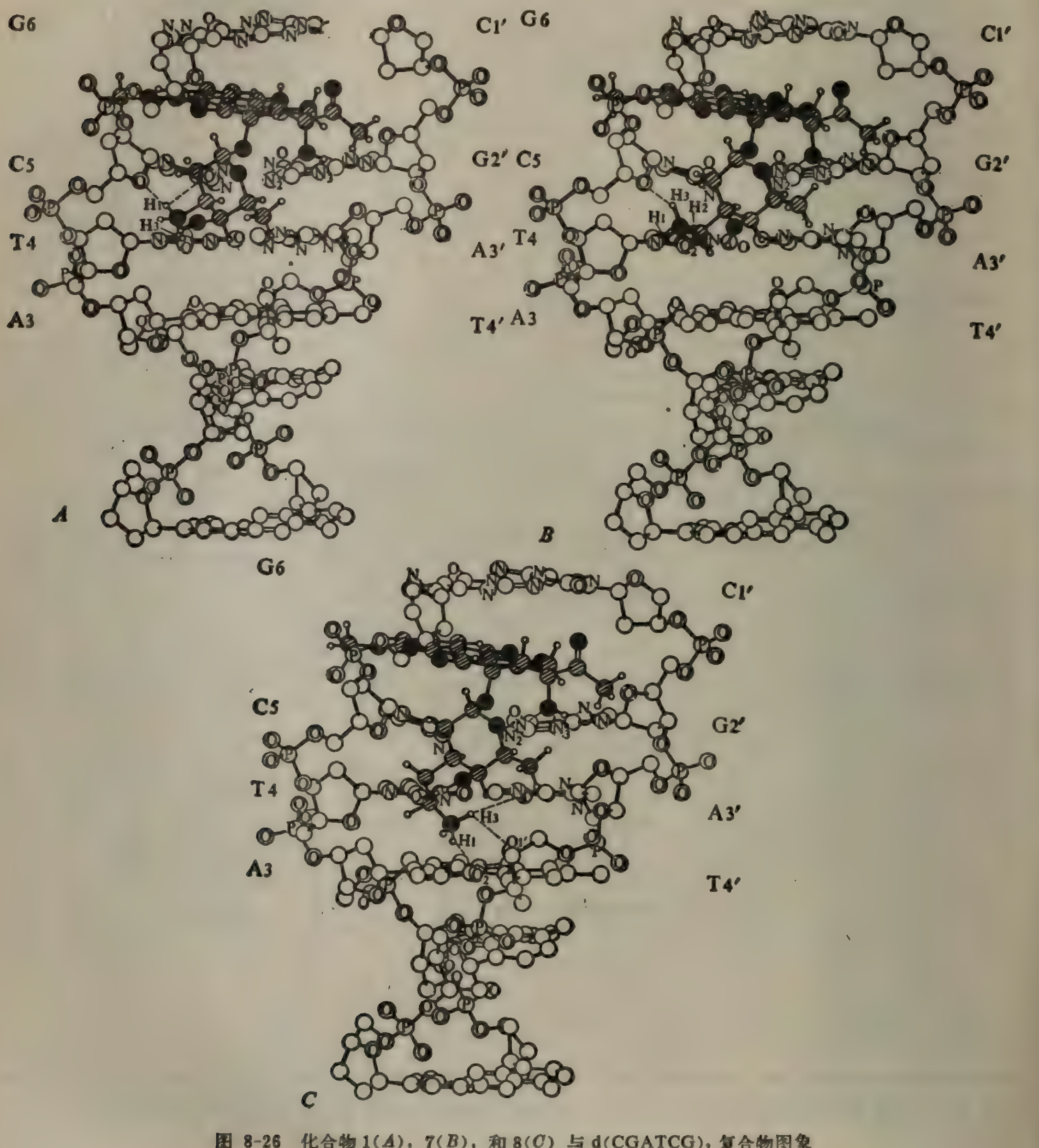


图 8-26 化合物 1(*A*)，7(*B*)，和 8(*C*) 与 $d(CGATCG)_2$ 复合物图象

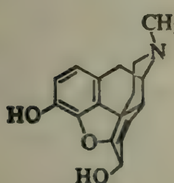
8.2.6 作用于中枢神经系统药物药效构象的研究

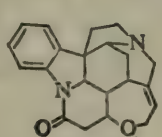
不同结构的药物产生相同的药理作用，如作用于中枢神经系统的药物，它们与受体作用有没

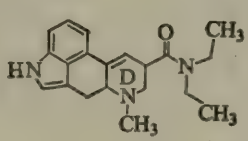
有共同的药效基团亦即一个共同的必要构象呢？

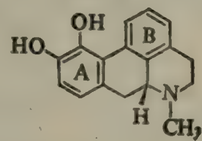
长期以来，人们认识到大多数作用于中枢神经系统（CNS）的药物分子中必定具备一个芳环和一个氮原子。近年来 Andrews 等进行的构象研究，证实了这二个基团独特的图象排列是各类 CNS 药物活性的最低要求[39-41]。

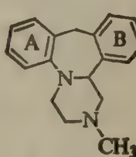
他们在进行 CNS 药物共同的结构模型(common structure model)研究中，将常用的 CNS 药物中的各种不同作用类型的分子共 14 个综合在一起，这些药物具有不同的结构特征。吗啡(1)（镇痛药）、马钱子碱(2)（中枢兴奋药）、二乙基麦角酰胺(LSD)(3)（致幻药）、阿扑吗啡(4)（镇痛药）、米安舍林(5)（抗抑郁药），这五个分子属于半刚性化合物(semirid compound)；苯巴比妥(6)（催眠药）、可氯宁(7)（肾上腺素 α 受体激动剂）、安定(8)（抗焦虑药）、荷包牡丹碱(9)（GABA 拮抗药）、苯妥因(10)（抗惊药），这五个化合物都在两个芳环间存在可自由旋转的扭角；苯丙胺(11)（兴奋药）、丙咪嗪(12)（抗抑郁药）、氯丙嗪(13)（抗精神病药）、卡马特灵(14)（抗胆碱药）这四个化合物都具有两个以上自由旋转的侧链。

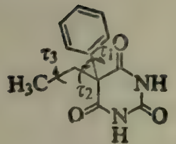

1

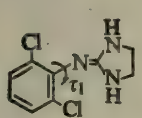

2

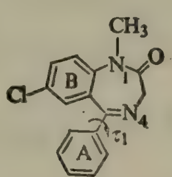

3

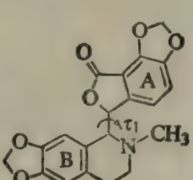

4

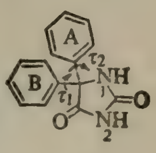

5

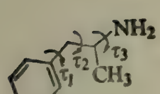

6

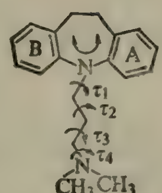

7

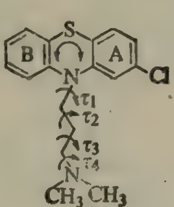

8

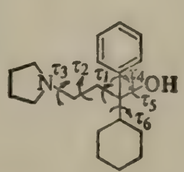

9


10


11


12


13


14

Andrews 采用分子力学计算方法，称之为经典势能计算法（classical potential energy）对上述 14 个分子进行了研究。他们最后得出 CNS 药物四点式作用的共同构象特征的论点。见图 8-27。

图 8-27 显示 CNS 药物中，处于芳环平面正上方和正下方各有一个受体作用点 R_1，其坐标(nm)为(0,0.35,0)，R_2，其坐标为(0,−0.35,0)，氮原子，坐标为(0.48,−0.03,0.14)，R_3，其坐

标为(0.63,0.10,0)。图中还表示 R_1 和 R_2 作 用点分别表示该药物与受体的疏水结合 (hydrophobic bonding)，R_3 表示与受体的氢键结合。Andrews 称之为共同模式(the Common Model)。

下面，让我们简要地分析一下作者所计算的 14 个化合物的具体结果。

对于 1—5 半刚性分子优化的势能计算相对较为简便，它们的六种可能的构象计算值列于表8-6，并运用计算机图形计算 MORPHEUS 系统进行比较分析，对应的空间分子图画于图 8-28。

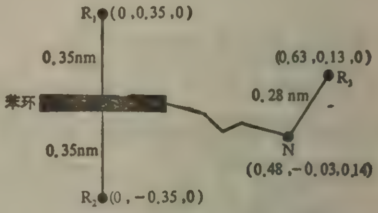


图 8-27 CNS 药物四点式作用构象示意图

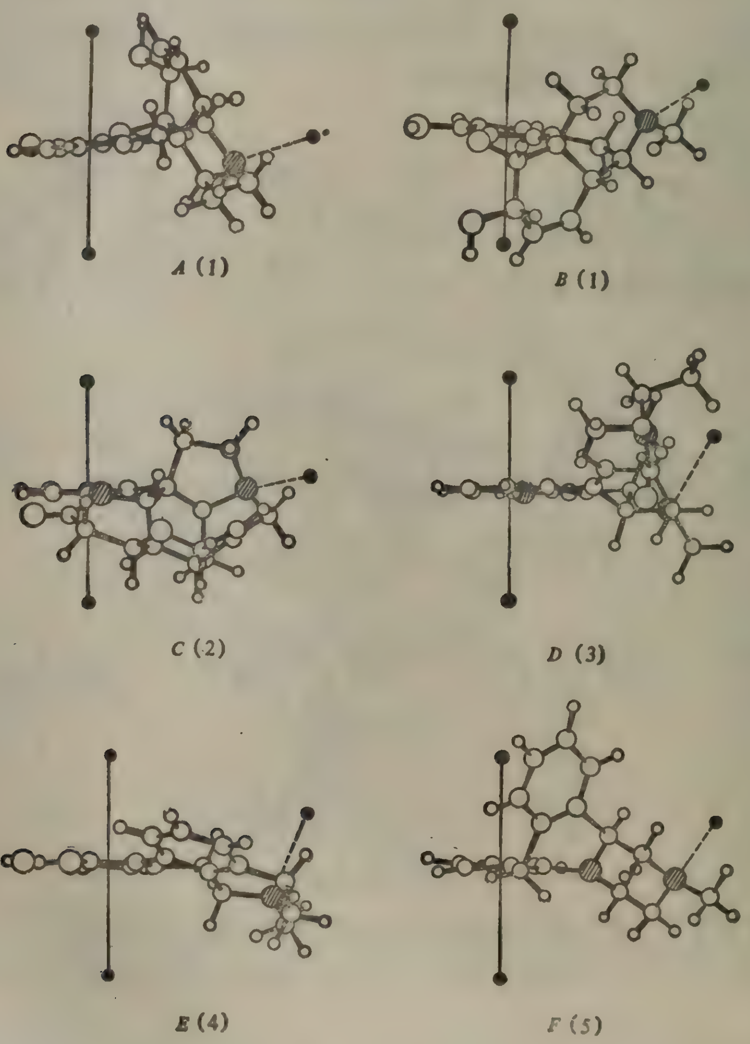


图 8-28 对应表 8-6 的六种构象的空间分子图

表 8-6 半刚性分子与共同模式的吻合

分　　子	构　　象	代　号	能量(kcal/mol)	距离(nm)*
吗　　啡(1)	N-R_3为轴	*A*	0.0	0.030
	N-R_3为直径	*B*	2.1	0.042
马钱子碱(2)		*C*	0.0	0.043
LSD(3)	N-R_3为轴	*D*	0.0	0.041
阿扑吗啡(4)	N-R_3为轴	*E*	0.0	0.062
米安舍林(5)	N-R_3为轴	*F*	0.0	0.033

* 最好吻合距离，用均方根测量(RMS)

对于 6—9 主要由一个旋转扭角确定构象的四个分子，进行扭角-势能的曲线计算，化合物 6 的计算结果见图 8-29，四个化合物的能量极小对应的构象见表 8-7，并同样画出四个化合物分子的五种构象的分子透视图——图 8-30。

对于有多个扭角的 10—14 这五个化合物，作者区别不同情况分别一一作了构象能量计算，图 8-31 显示了化合物 11 τ_1 和 τ_2旋转时所画出的势能构象图与最小距离图，可见 $\tau_1=110°$，$\tau_2=180°$为对应的优势构象。以上五个化合物的分子透视图列于图 8-32。

若将上述图 8-28、30(336页)、32(337页)中，除了 30*a* 和 32*d*外的这 14 个分子构象重迭，令人兴奋地发现，它们的芳环重迭外，假设的受体作用点 R_3 也重迭在一处(见338页图 8-33)。

上述研究充分表明 CNS 药物都有共同的结构基础，可能次级的结合部位才决定该药物的具体作用类型。这一类共同结构模型的研究仍在不断发展之中[42]。

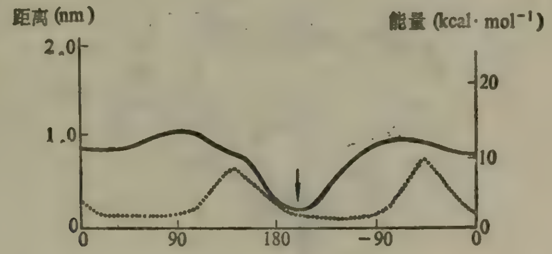


图 8-29 化合物 6 (苯巴比妥)的主要扭角与受体作用点之间的势能(——)和距离(……)曲线，表明一160°能量最低

8.3 药物的受体

8.3.1 受体学说在药物研究中的重要性

我们知道，各种类型的药物，可防治不同的疾病。但为什么这些化合物会产生疗效呢？结构截然不同的药物，为什么又对同样的疾病有疗效呢？这是药物学家早在上世纪初就开始思索的

表 8-7　分子6—9 的最低能量和最适距离与共同模式的吻合

构　象	代 号	扭 角	能量(kcal/mol)	距离(RMS) (nm)
苯巴比妥(6)	*A*	−160	0.0	0.03
可 氯 宁(7)	*B*	70	0.7	0.04
安定(8)　环A-N_1 S isomer	*E*	−100	0.1	0.10
安定(8)　环A-N_1 R		100	0.0	0.08
安定(8)　环A-N_1 三角形		−90	0.1	0.06
安定(8)　环B-N_4 S	*D*	−90	0.0	0.10
安定(8)　环B-N_4 R		−90	0.0	0.09
安定(8)　环B-N_4 三角形		−90	0.0	0.09
荷包牡丹碱(9)	*C*	−100	0.4	0.07

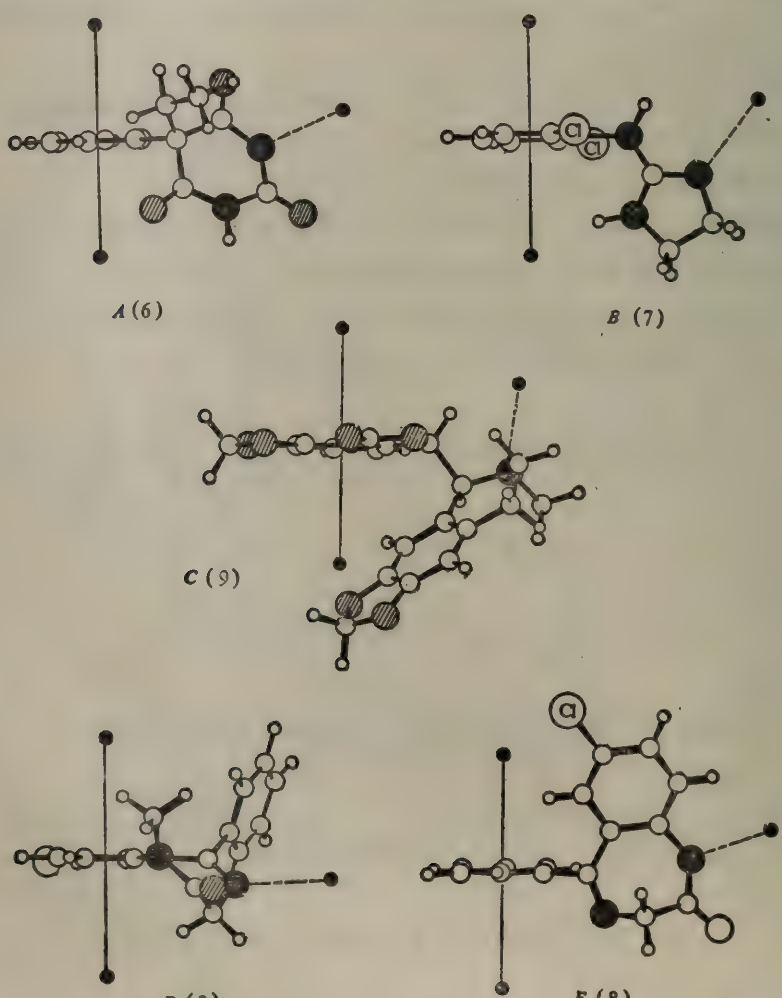


图 8-30　对应表 8-7 中 4 个分子 5 种构象的空间分子图

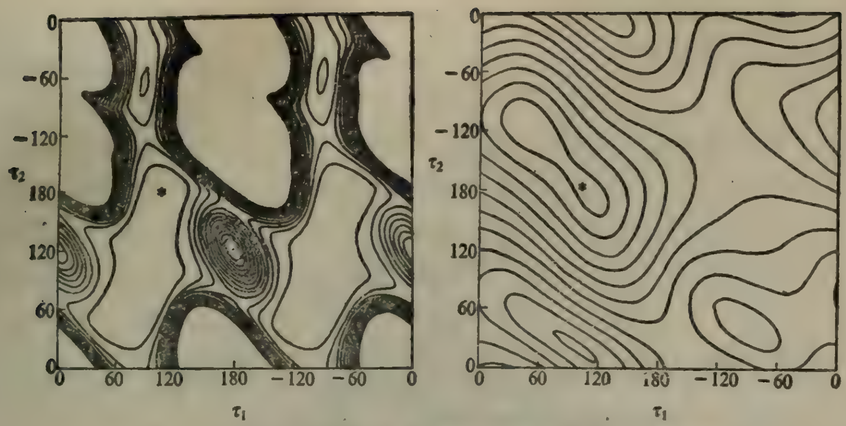
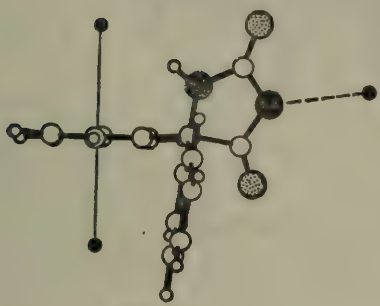
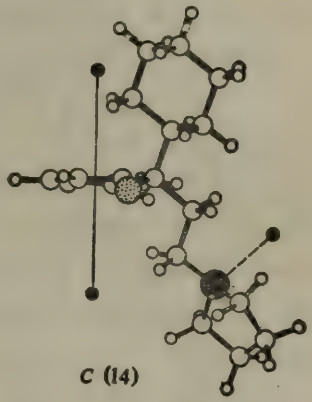


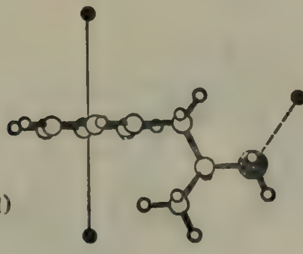
图 8-31 分子(11)扭角 $\tau_1\tau_2$ 旋转对应的势能图(*A*)和距离图(*B*)



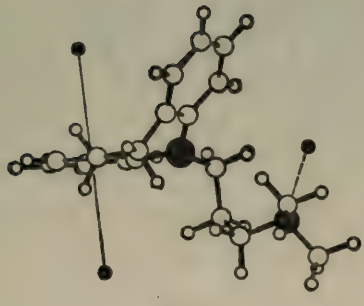
A (10)



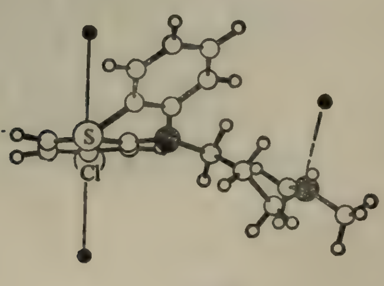
C (14)



B (11)



D (12)



E (13)

图 8-32 分子(10)—(14)空间分子图

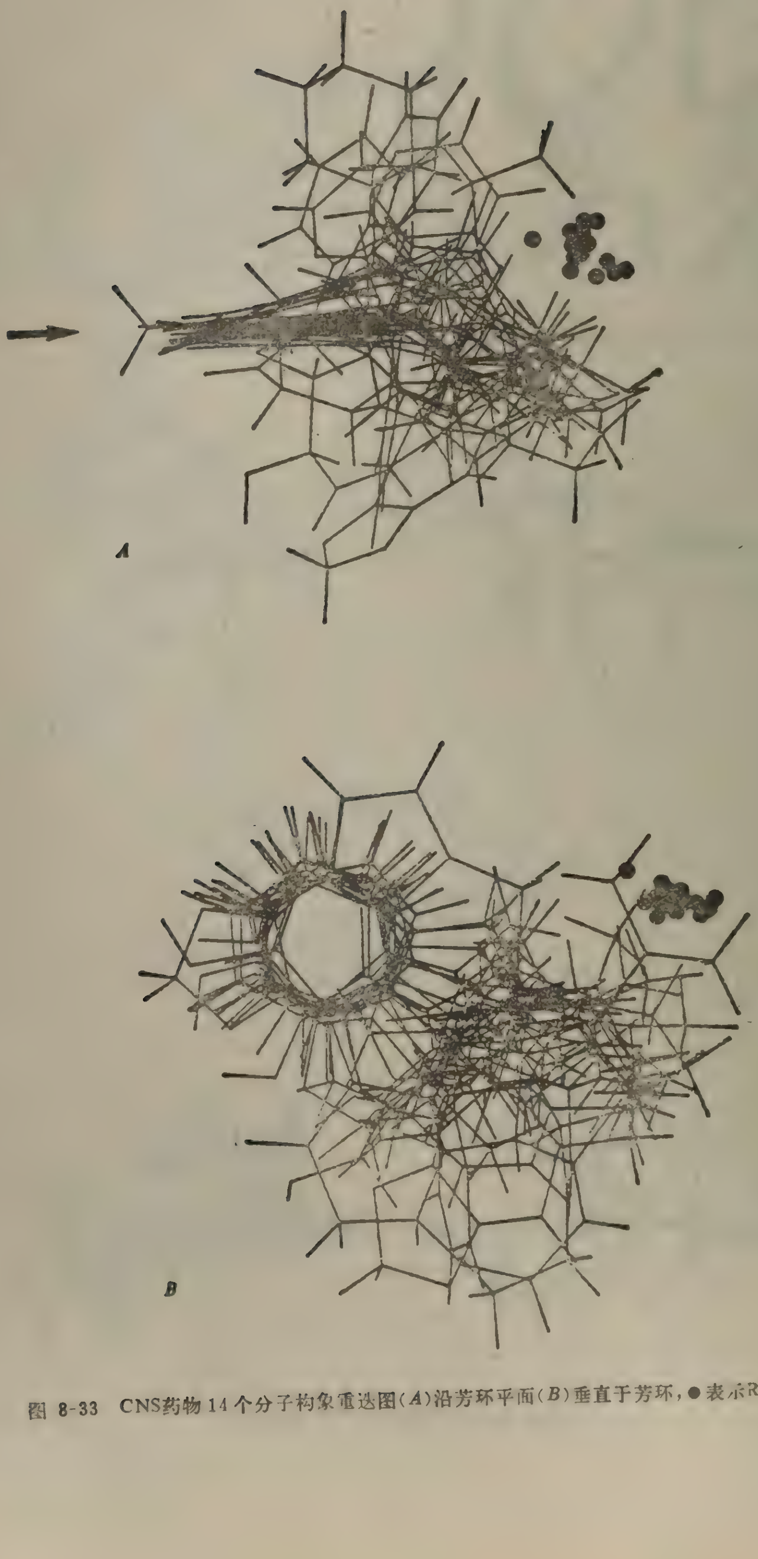


图 8-33 CNS药物 14 个分子构象重迭图(*A*)沿芳环平面(*B*)垂直于芳环，● 表示R

问题。1905年，Langley 就提出了受体(receptor)概念。所谓受体，是指生物体细胞的一种组分，当药物与这组分结合后，促使产生一系列变化，最终导致产生药理效应，如药物结合在细菌细胞，抑制细菌的生长；药物结合在血管，促使其舒张而降低血压。

早年，受体只是科学家的一种设想，人们一直沿用这一名称，并试图推测受体的形象。只有在分子生物学取得重大进展的今天，以及应用 X-射线衍射、NMR测定以及放射性示踪等若干新技术后，才开始跨上了研究受体本质，以及受体与小分子化合物相互作用的道路。

分子生物学研究揭示了酶与基质间的相互关系。它提示我们，药物的受体也可能类似酶。受体实际上有一个弹性的三维构象，其立体结构应当与药物的结构互补。有时当药物作用时，受体构象先发生变化，然后与药物的结构互补。两者互相结合而生成复合物以后，便将整个蛋白或邻近的蛋白激活，从而产生相应的生物活性。

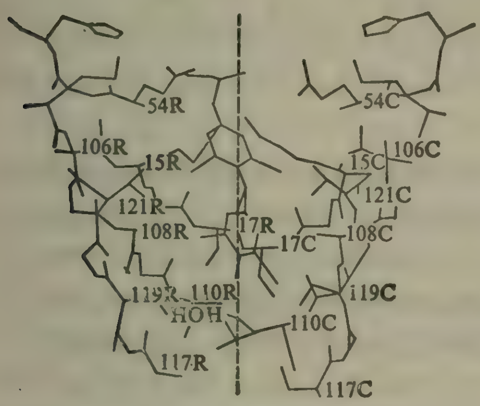


图 8-34 甲状腺素与蛋白质结合示意

在许多生化反应中，酶起着催化的功能。许多药物的药效，是由于与酶上某个部位的结合，从而抑制了酶的正常催化功能。这样，这些药物的受体是酶上的部位。研究复杂的酶的三维结构表明，赖氨酸、精氨酸、组氨酸、丝氨酸、苏氨酸、酪氨酸、色氨酸、谷氨酸、天门冬氨酸等氨基酸的极性基团，绝大多数都在蛋白分子的外表部分。药物与酶的结合，正是通过与氨基酸上取代基的作用力。图 8-34 显示了甲状腺素与蛋白质上各种氨基酸上基团的结合。

除了蛋白是药物受体以外，核酸也是重要的药物作用部位。例如蒽环素族(Anthracyclines)抗生素插入在双螺旋核酸的二个邻近碱基间，使核酸的构象产生变异[43]。从而干扰其样板作用。其他如抗血吸虫药物甲硫蒽酮、吖啶类抗疟药、阿霉素等抗肿瘤药物也嵌在 DNA 的两个核苷酸之间。

在药物与受体生成复合物的过程中，两者的一些官能团间产生作用，这些作用不外乎有以下不同的形式。

(1) 共价键。多数药物与受体间作用是可逆的。当细胞外液的药物浓度降低时，这复合物又解离。但少数药物如烷化剂、酰化剂等，可与受体产生共价键结合。如抗肾上腺素药苯氧苄胺的 $N-CH_2CH_2Cl$ 基团，可与酶上的巯基等活性基团以共价键结合。

(2) 离子键。药物分子内的羧基或磺酰胺在生理酸碱度条件下(pH=7.4)可电离为阴离子，与受体的阳离子产生静电结合，另外，药物分子的脂肪胺可质子化成为阳离子，季铵也以阳离子形式存在，它们可与受体上的阴离子生成离子键。

(3) 氢键。羧基、氨基、羟基、亚氨基等均可作为氢键的接受体或给予体。

(4) Van der Waals 引力。虽然其作用力很弱，但许多基团的引力的协同作用亦不可忽视。

(5) 疏水性结合。药物的非极性部分(如烃基)与受体的非极性部分，可挤去外围的水分子，产生引力。

(6) 电荷转移复合。羰基或芳环等是 π 键或大 π 键，电子易于流动，药物与受体间这种基团可能产生电荷转移复合。

尽管已经分离出不少神经递质和激素的受体，但都是分子量达好几万的生物大分子，况且蛋白质本身又很容易变性，这就很难研究与药物相结合的受体的基团所在。目前的研究方法，先从药物本身的结构开始。药物既与受体相结合，两者结构自应有互补性。则可从药物分子的构象、电荷分布、空间的静电势场、化学键活性、疏水性能等角度，可以反推与其作用的受体结构。本章第二节阐述的药物分子的构象，实际上是研究药物受体的需要，研究构象的目的是探讨受体的性质和结构。当然，X 衍射足以测定药物的三维结构，若带有不对称中心，所带基团的空间排列即绝对构型也可测定，晶体状态代表最低能量时的分子构象，但不一定是“活性构象”，因为晶体结构也可能转化为另一种构象，再与受体结合。NMR获得的数据，包括偶合常数与化学位移，亦可推算溶液中不同构象的存在比例，它以溶液中存在的构象代表体液环境的药物，当比晶体结构更为符合实际。

量子化学分子轨道法计算可方便地计算药物分子的能量，各原子的电荷密度、空间静电势、化学键键级等，正如第二节所述，用量子化学计算不同结构的数种药物，从其中几个能量较低的构象中，可取其共同性的构象，可能就是药效构象(pharmacophore)。经量子化学计算所得的分子内电荷密度分布，也可反推受体各相互作用点的性质，电荷密度较高的部位，可作用于受体的正电中心，反之电荷密度较低的部位，则可作用于受体的负电中心。福井谦一建立的量子化学前沿轨道指数，亲电、亲核前沿轨道电荷密度f^E、f^N；亲电、亲核、亲自由基超离域度 S_r^E、S_r^N、S_r^R，这些指数在研究药物与受体间的作用上起了很重要的作用。分子静电势又可描绘分子周围电荷作用，其正负电势部分可清楚地分别发生亲电或亲核的进攻。此外可从计算所得的最高占据轨道能量(E_{HOMO})和最低空轨道能量(E_{LUMO})的大小分析，当药物与受体相互作用时，其电子跃迁能力的大小。这些均对受体的研究产生重大的推动作用。

研究体内的生理活性物质或药物与受体怎样结合，不仅是为了揭示自然界的奥秘，而且更是为了合理地开发和设计新药。在研究探索了药物与受体怎样作用以后，将使药物结构的设计更趋合理。当前有关受体的知识还是比较肤浅，根据受体的研究而设计成功的新药为数很少，然而受体的研究及其应用，必将使药物设计进入一个新的天地。从这个意义上讲，受体学说是药物学领域的一个极有希望的发展方向。

关于受体的研究 Richards[1]在他的论著中给予很重要的位置。嵇汝运多次发表综述，阐述了受体研究的重要性以及在新药设计中的应用[2,11,44]。本节主要介绍量子化学在受体研究中的一些直接应用，第二节的构象研究及随后的药物分子静电势与电荷密度，以及电荷转移，事实上都万变不离其宗，均与受体学说息息相关。

8.3.2 乙酰胆碱的量子药理学研究

乙酰胆碱与去甲肾上腺素分别是两类不同神经系统的神经冲动传递媒介，称为神经递质。

现已研究清楚，副交感神经受到兴奋时，其节前纤维的末梢释放乙酰胆碱(acetylcholine)，见图8-35，后者乃引起节后纤维的兴奋，于是节后纤维的末梢也释放乙酰胆碱。近年来，将不同递质的两类神经称为胆碱能神经与肾上腺素能神经。胆碱能神经是指以释放乙酰胆碱作为神经冲动传递媒介的神经纤维。



图 8-35　乙酰胆碱分子透视图

胆碱能神经包括：(1) 全部交感神经和副交感神经的节前纤维；(2) 全部副交感神经节后纤维；(3) 小部分交感神经节后纤维(如支配汗腺的纤维)；(4) 运动神经。这就意味着一个神经冲动从节前纤维传递到节后纤维，自神经纤维传递至骨骼肌或平滑肌，或自交感神经传递至汗腺等腺体，都须依赖乙酰胆碱的作用。它的作用范围如此广泛，可见它在神经药理中的地位是十分重要的。

乙酰胆碱在药理学中的地位理所当然地引起了量子药理学研究的重视。可以说迄今为止对乙酰胆碱的量子化学研究以及作用于胆碱能神经药物的研究是量子药理学中工作最多的一部分。几乎所有的适用于量子药理学研究的量子化学常用方法，如 EHMO、PCILO、CNDO/2、INDO，ab initio 等都在该领域内大显身手。详细的文献回顾至 1981 年可参见文献[7]。

Pullman 等首次使用从头计算法，采用 STO-3G 基组，用 Gaussian 70 程序对乙酰胆碱进行了构象优选计算。图 8-36 显示了乙酰胆碱分子的构象旋转图，可见影响构象的主要因素是 τ_1 和 τ_2 旋转。

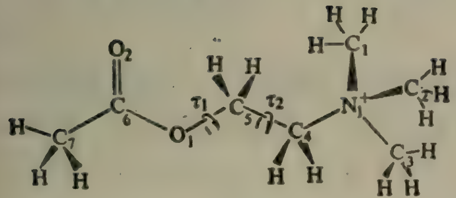


图 8-36　乙酰胆碱的分子旋转

$\tau_1=\tau(C_6-O_1-C_5-C_4)$

$\tau_2=\tau(O_1-C_5-C_4-N^+_1)$

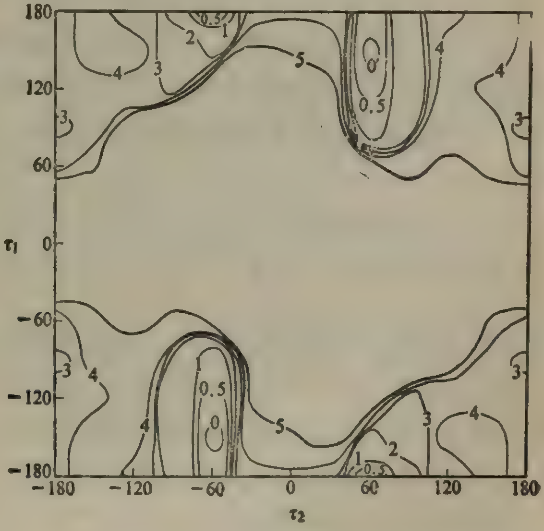


图 8-37　用 ab initio(STO-3G)获得的乙酰胆碱分子构象能图(kcal/mol)

图 8 - 37 显示了经 ab initio 计算得到的乙酰胆碱构象能图，表明 $\tau_1=150°$、$\tau_2=60°$时能量最低。

图 8 - 38 是用最低能量的构象计算所获的乙酰胆碱各原子净电荷值。

从图中可见，经从头计算法计算，乙酰胆碱分子中 O1 和 O2 原子各带有较多的电子。$-N(CH_3)_3$ 上的净电荷分布可以看出，尽管氮原子本身带有电子(−0.157)，然而，这个基团的正电荷被三个甲基所分散，9 个氢原子缺少 1.098 个电子，这表明$-N(CH_3)_3$对受体作用显示强烈的亲电性。

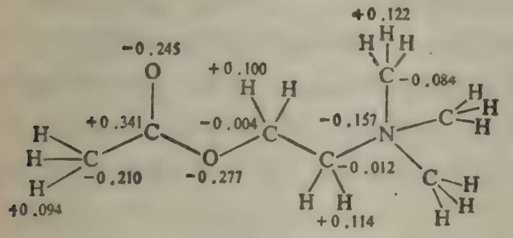


图 8-38 从头计算法所得乙酰胆碱分子各原子净电荷值（H 原子取相同原子的平均值）

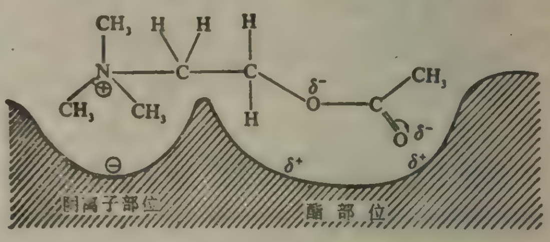


图 8-39 推测的乙酰胆碱-受体复合物的典型图象

Richards 形象地画出了乙酰胆碱受体作用机制图。图 8 - 39 是经过量子药理研究推测的乙酰胆碱-受体复合物的典型图象。

这种经过量子药理学研究推测药物受体，是量子药理学主要内容，所画的受体制图迄今仍被广泛地采用。

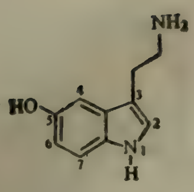
8.3.3 5-羟色胺受体

有些化合物服用后，会使人产生幻想、幻觉、幻视、幻听，近似精神病患者的精神状态，这类化合物称致幻剂。致幻剂的量子化学计算是量子药理学的最早研究成果之一。[45]致幻剂产生如此奇幻作用的原因，是由于其作用在中枢神经递质5-羟色胺受体的缘故，从而干扰了正常的神经递质的作用。

5-羟色胺也称为 Serotonin 或 5 -HT。人们熟知的致幻剂有二乙基麦角酰胺(LSD)、墨西哥致幻蕈素(psilocin)、仙人球毒碱(mescaline)、1-甲基-2-(2,5-二甲氧基-4-甲基)苯乙胺(DOM)。

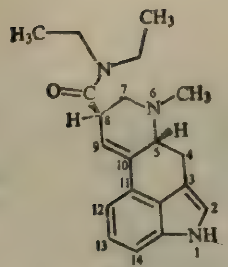
分析这些致幻剂，既然作用于 5 -羟色胺受体，则它们的药效基团应与 5 -羟色胺有着相似排列。用量子化学 INDO 法计算並结合 X 衍射及 NMR 测定，获得了预期的结果[40]。图 8 - 40 显示了墨西哥致幻蕈素用量子化学构象势能研究的 τ_1、τ_2 旋转示意图。经优选后，$\tau_1=90°$，$\tau_2=180°$。图 8 - 41 和图 8 - 42 分别是用 INDO 方法计算所获的构象势能图[46]。

从图 8 - 43 可以全面看出致幻剂化合物 *B*、*C*、*D*、*E* 的构象与 5 -羟色胺构象其关键基团有着相同的排列。图 8 - 44 显示了 DOM 与 5 -羟色胺药效构象的重合。黑色键是 5 -羟色胺分子，可见其氮原子、氧原子重叠、吻合。



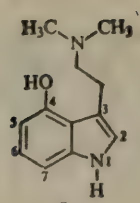
A

5-羟色胺



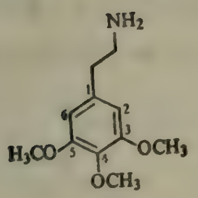
B

二乙基麦角酰胺(LSD)



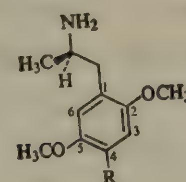
C

墨西哥致幻蕈素



D

仙人球毒碱



E

$R=CH_3$

DOM

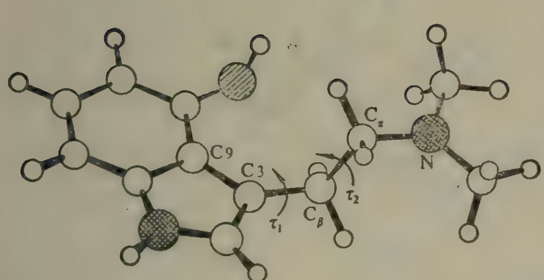


图 8-40 墨西哥致幻蕈素构象旋转图

$\tau_1=\tau(C_\alpha-C_\beta-C_3-C_9)$

$\tau_2=\tau(N-C_\alpha-C_\beta-C_3)$

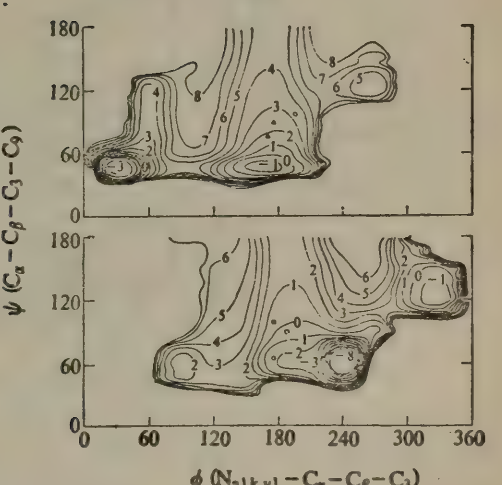


图 8-41 用 INDO 法计算所获的墨西哥致幻蕈素的构象势能图 ϕ 对应 τ_2 旋转，ψ 对应 τ_1 旋转

这样构象分析的目的是给致幻剂 5 -羟色胺受体的研究以一有力的佐证。

关于 5 -羟色胺受体的量子化学研究现仍在不断探索之中。Mazurek 与 Weinstein 等对 5 -羟色胺类似物的药物-受体相互作用进行了实验测定和量子化学理论研究[47]。在研究的 5 -羟色胺(I)和两个类似物 FHATHBIN(II)、FAEFHI(III)中，(II)分子中，由于六元环的结构，使其

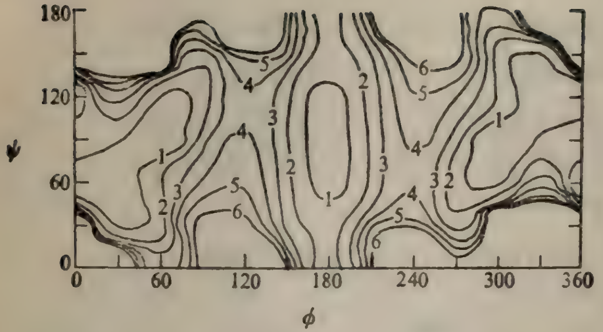


图 8-42　用 INDO 法计算所获的仙人球毒碱的构象势能图

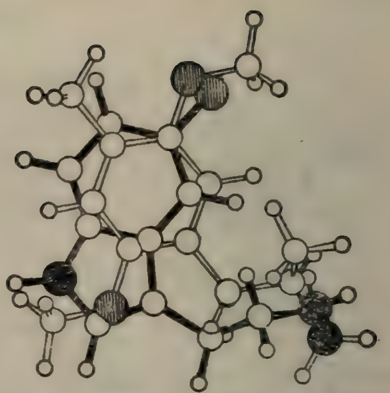


图 8-44　DOM 与 5-羟色胺的药效构象重合

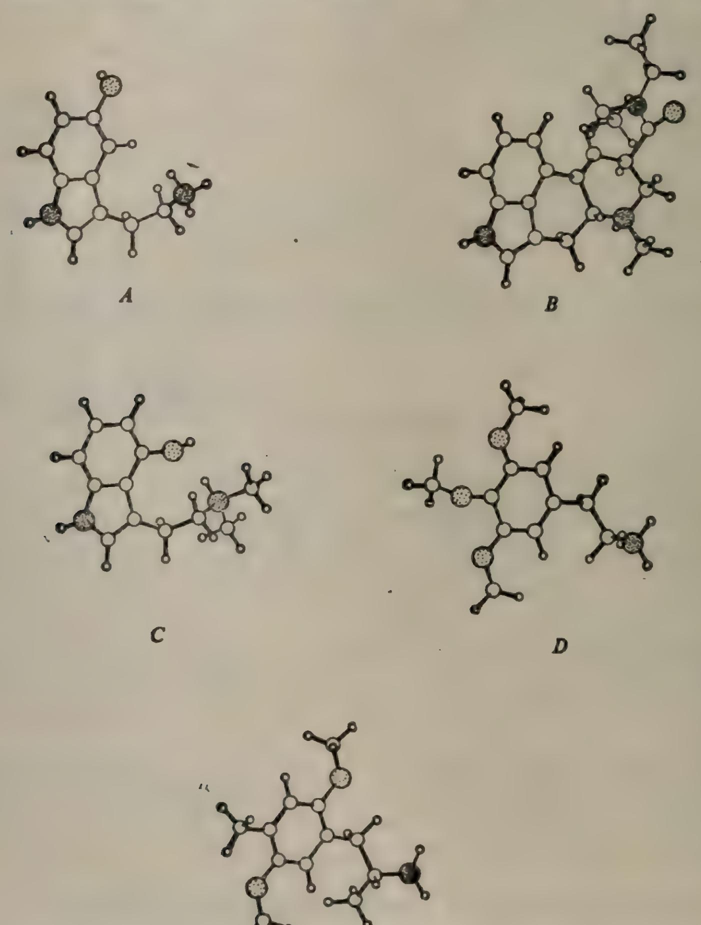
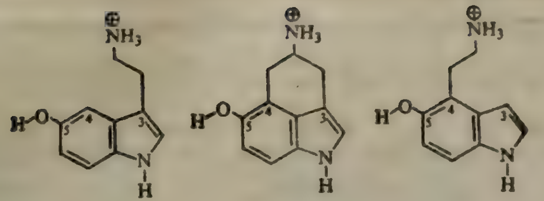


图 8-43　5-羟色胺(*A*)、二乙基麦角酰胺(*B*)、墨西哥致幻蕈素(*C*)、仙人球毒碱(*D*)、DOM(*E*)的药效构象



构象与5-羟色胺的延伸型构象吻合，因而其与5-羟色胺受体的作用活性也相近。分子(III)与(I)相比，它们是异构体，所不同的仅是乙胺基侧链从3位移至4位。看起来，(III)也有类似的延伸型构象，当然也应有相近的生物活性。然而经实验测定，(III)的生物活性仅为5-羟色胺的0.5%～0.6%。这种特异现象引起量子化学家的极大兴趣。作者采用从头计算法，使用Gaussian 80程序，对图8-45所示的分子(III)的五种不同的分子构象进行计算，为了便于采用不同基组计算结果(如STO-3G、3-21G)，使用分子片断模型进行计算，以及运用Pople的方法，对整个分子局部采用3-21G而其余都用极小基组(STO-3G)的技术[48]，所获结果列于表8-8。

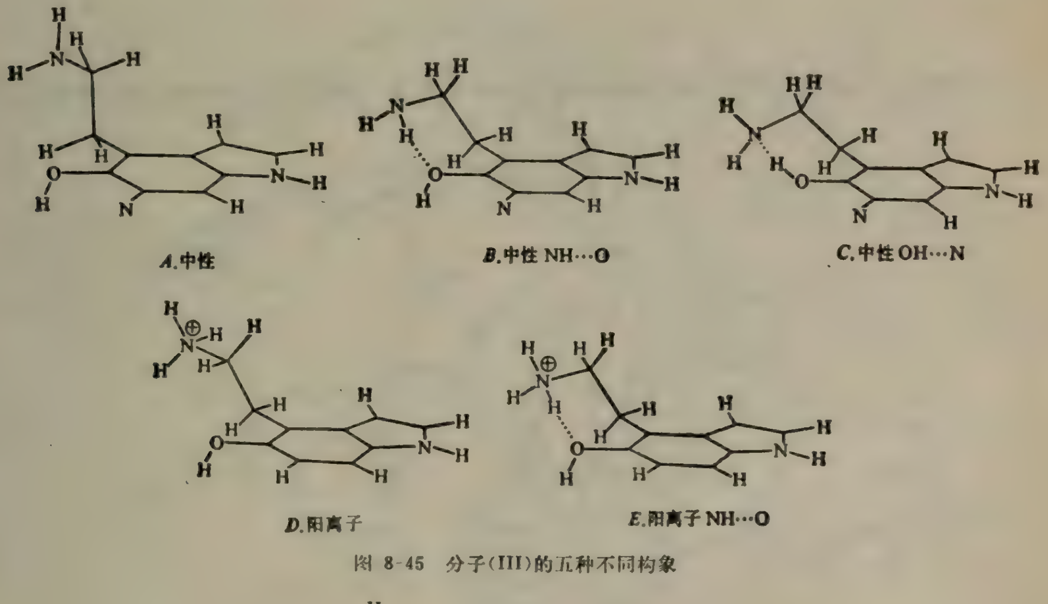


图 8-45　分子(III)的五种不同构象

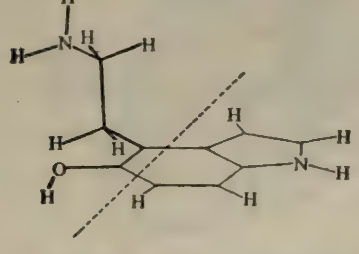


图 8-46　分子(III)的模型片断

表 8-8 FAEFHI 和模型分子的稳定能(kcal/mol)与总能量(Hartree)

基 组	3-21G		STO-3G		STO-3G//3-21G	
化合物	Model		Model		FAEFHI	
	总能量	稳定能	总能量	稳定能	总能量	稳定能
A	−284.378115		−282.360895		−562.330375	
B	−284.380154	1.3	−282.363027	1.3	−562.333239	1.8
C	−284.396814	11.7	−282.375686	9.3	−562.337201	4.3
D	−284.774539		−282.805348		−562.782152	
E	−284.791184	10.5	−282.832591	17.1	−562.810958	18.1

注：第三列 STO-3G//3-21G 表示分子模型片断采用 3-21G 而其余原子用 STO-3G 基组计算

从上述结果可知，FAEFHI 的乙胺基侧链与 5 位羟基形成分子内氢键的构象能量降低，如表中第一列，形成OH……N 氢键的构象(*C*)，其总能量比不形成内氢键的构象(*A*)降低 11.7 kcal/mol。因此分子内氢键的形成使得 5-羟色胺受体要求匹配的延伸型构象所占比例极少，故而活性大大降低。

8.3.4 酶 抑 制 剂

不少药物的作用部位在酶。激动剂(如神经递质或激素)与酶上受体结合后，便将整个蛋白

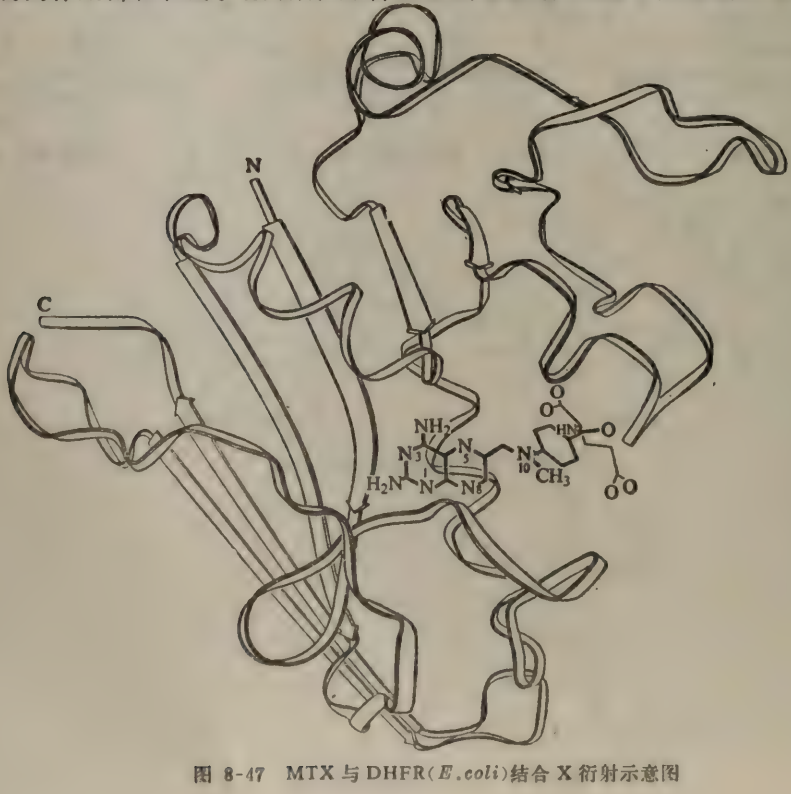
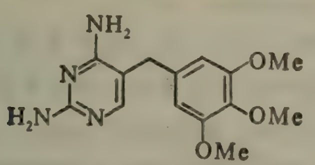


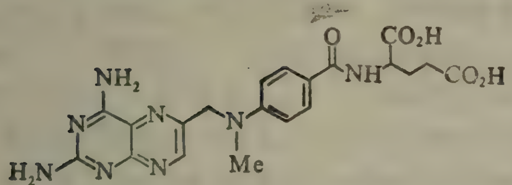
图 8-47 MTX 与 DHFR(*E.coli*)结合 X 衍射示意图

或邻近的蛋白激活,产生作用。而多数药物是酶的抑制剂。已经合成的有效酶抑制剂为甲氧苄氨嘧啶(trimethoprim,TMP)与氨甲喋呤(methotrexate,MTX),它们是二氢叶酸还原酶(DHFR)的抑制剂,干扰二氢叶酸还原为四氢叶酸,从而阻止细菌合成核酸所需的胸腺嘧啶核苷,细菌生成乃受抑制。

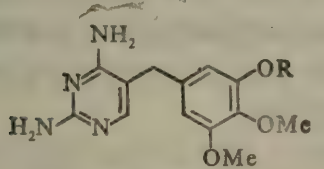
DHFR 酶的氨基酸顺序业已阐明,其与 MTX 结合的整体构象已经由 X 衍射研究而获得。图 8-47 表示 MTX 与从大肠杆菌(*E.coli*)中分离的 DHFR 结合的 X 衍射结晶示意图[49]。图中可见,MTX 的杂环部分结合在酶的疏水性囊中,酶上第 6 与第 7 位二个丙氨酸所成肽键与环平行,可能与药物生成电荷转移复合物。第 113 位的苏氨酸的羟基可能与药物第 2 位的氨基产生氢键,药物第 1 位的氮与酶上第 27 位天冬氨酸的羧基相近,可能以离子键相结合。药物侧链上的对氨基苯甲酰基结合在酶上的第二个疏水性囊。第 28 位与 54 位的亮氨酸、第 50 位与94 位的异亮氨酸以及第 49 位丝氨酸与药物侧链氮上的取代甲基以 Van der Waals 引力相作用。药物分子 α 位的羧基与酶上第 57 位精氨酸的胍基生成离子键,而 γ-羧基则在酶的第 32 位赖氨酸与第 52 位精氨酸的碱性侧链附近。



1.TMP



2.MTX



3.R= $(CH_2)_5COOH$

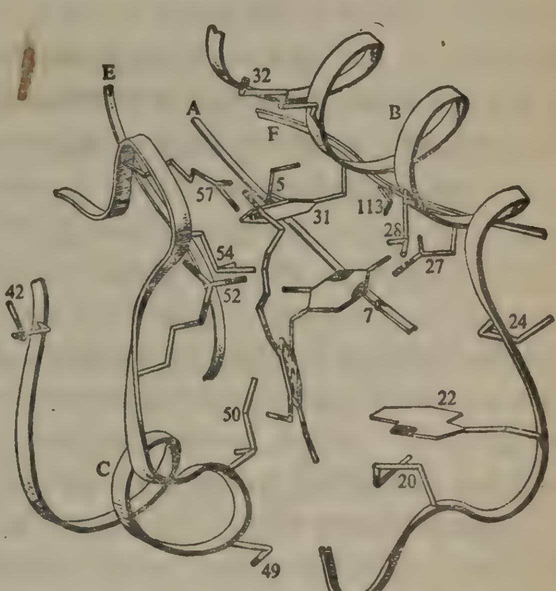


图 8-48 化合物(3)与 DHFR(*E. coli*)结合 X 衍射示意图

TMP 分子上并无羧基,这样就会与前述碱性氨基酸的结合不强,因而设计将苯环上的一个甲氧基改变为羧戊氧基(如 3),羧基便可与酶上第 32、52、57 位的碱性氨基酸作用,从而增强抑制酶的作用。图 8-48 是化合物(3)与 DHFR(*E. coli*)结合的 X 衍射示意图[50]。

经过深入的研究,发现 TMP 具有很突出的选择性,它对细菌抑制作用强,而对人体毒性低。从 TMP 对大肠杆菌(*E. coli*)DHFR 和对正常人体脾淋巴细胞 DHFR 的抑制作用分析,

正常人体脾淋巴细胞(Human WIL$_2$) DHFR 对 TMP 的敏感性要比 *E. coli* DHFR 低 66 000 倍[51]。由于各种生物来源的 DHFR 在结构上的差异，使得它们对于同一种抑制剂会表现出不同程度的敏感性，这就为研究药物的选择性作用提供了很好的课题。同时，对各种生物来源的 DHFR 与抑制剂间相互作用的研究所得的信息，将对设计具有更高选择性作用的抗菌和抗癌药物有指导性作用。

对不同生物来源 DHFR 的抑制作用研究，亦称离体酶研究，是近年来药物研究发展的重大成果之一。Hansch 等[52]首先研究了一系列 TMP 的衍生物——5-取代苄基-2，4-二氨基嘧啶类化合物的结构与各种生物来源的 DHFR 抑制作用的关系。在进行一系列连续的研究工作之后，导得了许多定量构效关系(QSAR)的相关方程式。

在对代表与微生物作用的 *E. coli* DHFR 的抑制作用的研究中，导得 QSAR 相关方程[53]

$$\begin{aligned}\log 1/K_{i_{app}} =& 0.43(\pm 0.015)\pi_{3,4,5} + 1.23(\pm 0.23)MR'_{3,5} \\ &+ 0.80(\pm 0.27)MR'_4 - 0.88(\pm 0.26)\log(\beta \cdot 10^{\pi_{3,4,5}} + 1) \\ &+ 0.45(\pm 0.28)\sigma_R^- + 6.23(\pm 0.13)\end{aligned} \tag{8.5}$$

$$n = 43;\ r = 0.923;\ s = 0.263;\ \pi_0 = 0.64;\ \log\beta = -0.67$$

式中 $MR'_{3,5}$ 是苄基 3,5 位上取代基 MR（克分子折射率）值的 1/10，当取代基 MR 值大于 0.79(亦即$-OCH_3$的 MR 值)时仍按 0.79 计算，MR'_4 代表 4 位取代基的克分子折射率值，它的处理同 $MR'_{3,5}$。$\pi_{3,4,5}$系 3,4,5 位取代基 π 值(Hansch-Fujita's 疏水常数)之和。

在对表示人体和动物毒性的研究中，利用正常人体脾淋巴细胞中分离出的 DHFR (Human WIL 2)测定 TMP 衍生物的抑制活性，并导出 QSAR 方程[54]

$$\begin{aligned}\log 1/K_{i_{app}} =& 0.59(\pm 0.20)\pi_{3,5} - 0.63(\pm 0.59)\log(\beta \cdot 10^{\pi_{3,5}} + 1) \\ &+ 0.19(\pm 0.14)\pi_4 + 0.19(\pm 0.15)MR_3 \\ &+ 0.30(\pm 0.28)\sum\sigma + 4.03(\pm 0.13)\end{aligned} \tag{8.6}$$

$$n = 38;\ r = 0.879;\ s = 0.266;\ \log\beta = -0.82;\ \pi_0 = 1.94$$

比较式(8.5)、(8.6)，可知 TMP 类似物对人体来源的 DHFR 的抑制作用主要与取代基疏水常数有关，而对来自细菌的 DHFR 的抑制作用主要取决于取代基的 MR'。这不但完全解释了 TMP 为什么是一个作用强、毒性低、选择性作用好的药物，而且可以依照这两个方程式继续有目的地定向合成高选择性化合物，使其选择性接近甚至超过 TMP。

对离体酶的各种理论计算的方法都获得了很大的成功。TMP 类似物已进行三维受体模型研究(three-dimensinoal receptor model, TDRM 法)[55]，其中有嘧啶类类似物 25 个，三嗪类化合物 14 个，并配合以量子化学 CNDO/2 法计算，给 QSAR 研究的结果以理论上的说明。

Andrews 运用分子力学方法和量子化学计算对 DHFR 抑制剂 TMP 和 MTX 类似物共 11 个作了深入的研究[56]。他们使用 COMOL 的称为经典势能计算的分子力学程序和 Allinger 发展的 MM 2 分子力学程序[14,15]在 Cyber 73 型和 VAX 11/780 型计算机上完成。由于 MM 2 程序在计算含杂原子 π 体系时缺少必要的参数，作者还使用 PRI/MINDO/3[57]量子化学程序进行了参数的优化。TMP 的分子力学计算以 τ_1 和 τ_2 的旋转为主(图 8-49)，其所绘制的二维势能构象图见图 8-50。

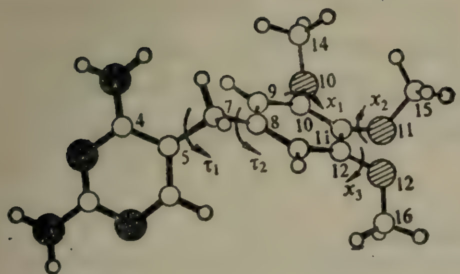


图 8-49 TMP分子 $\tau_1(C_4C_5C_7C_8)$, $\tau_2(C_5C_7C_8C_9)$, $\chi_1(C_9C_{10}O_{10}C_{14})$, $\chi_2(C_{10}C_{11}O_{11}C_{15})$, $\chi_3(C_{13}C_{12}O_{12}C_{16})$旋转示意图

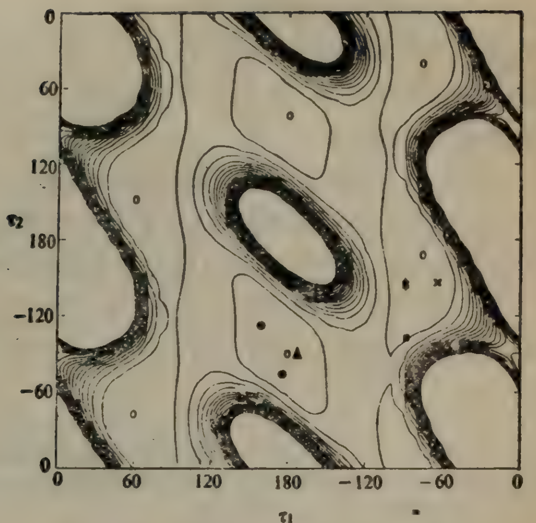


图 8-50 TMP 分子 $\tau_1\sim\tau_2$ 旋转势能图

上图中选取 $\chi_1=-80^\circ$, $\chi_2=90^\circ$, $\chi_3=80^\circ$，图中标注为零的极小能量区有 6 个。其中标有 * 的 $\tau_1\tau_2$ 对应角度为 TMP 单晶 X 衍射实验值，标有 • 处是 TMP 与 *E. coli* DHFR 结合的 X 衍射值，■ 位置是 NADPH-TMP 与鸡肝 DHFR 结合的 X 衍射实验值。实验结果表明 TMP 与细菌(*E. coli*)DHFR 结合时，其 $\tau_1=177^\circ$, $\tau_2=76^\circ$，苯环的位置与杂环的位置近于垂直。当与动物体(鸡肝)DHFR 结合时 $\tau_1=-85^\circ$, $\tau_2=102^\circ$，这两个芳香环的位置比较趋近于平行，这就显示了对不同来源 DHFR 的不同抑制活性[58]。而分子力学的计算结果所画出的能量极小区域又都包含了这些实验测定值，表明理论进行的构象计算与实验吻合。为进一步研究这些酶抑制剂与 DHFR 的作用，对 11 个化合物用 CNDO/2 法获得的波函数进行分子静电势计算[59]。图 8-51 显示了化合物 7(MTX)的分子静电势图。

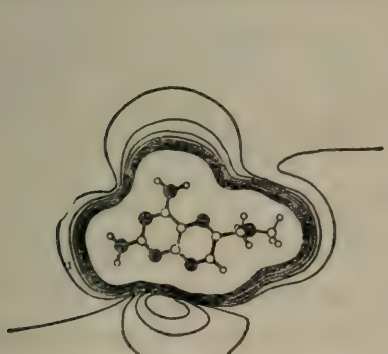


图 8-51 MTX 分子静电势图

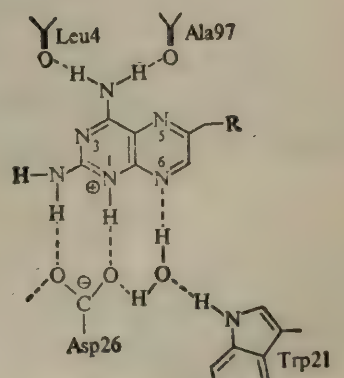
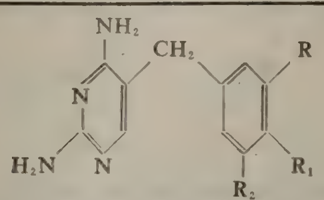


图 8-52 MTX 分子与(*L. casei*)DHFR 结合

分子静电势图显示在 MTX 分子中 N_1和 N_8下方是一个强的负电势区，它与 DHFR 结合时与酶(*E. coli* DHFR)上第 27 位天冬氨酸的羰基相近，有可能形成离子键，若与乳酸酪菌 (*L. casei*)DHFR 结合时，这个负电势区与酶上第 26 位天冬氨酸的羰基接近，可能形成 如图 8-52 所示的离子键结合。

依据分子力学计算的分子图形学(molecular graphics)研究还可应用于设计新的酶抑制剂。Czaplinsky等使用 CHEMMOD 分子力学软件包的分子图形学方法[60]，他们根据 TMP 衍生物中若有基团与酶上第 57 位精氨酸的胍基生成离子键，可增强抑制活性的分析，进行了分子模型的分子图形计算，表明引进如表 8-9 所示基团可增加抑制酶的活性。经有机化学合成后并进行抑制试验，完全证实理论预言的正确，于是一个新的强效酶抑制剂 K-130 经过理论设计 而最终问世。

表 8-9 TMP 及类似物结构及活性值



				E. coli		*M. lufu*	
				I_{50}	MIC	I_{50}	MIC
	R	R_1	R_2				
TMP	OCH_3	OCH_3	OCH_3	0.0055	1.5	0.07	>110
Brodimoprim	OCH_3	Br	OCH_3	0.0024	1.0	0.06	26.5
K-96	OCH_3	*1)	OCH_3	0.0153	11.3	0.19	30.7
K-107	OCH_3	*2)	OCH_3	0.0036	22.5	0.06	1.8
K-130	OCH_3	*3)	OCH_3	0.0020	56.0	0.01	1.0

*1) *2) *3)

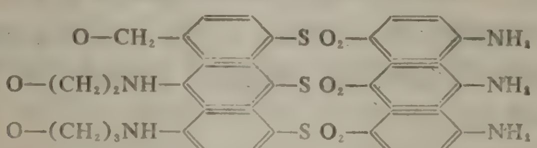


图 8-53 显示了 K-130 与 *E. coli* DHFR 结合示意图，图上清晰地可见，K-130 中 R_1 取代基的—SO_2—与酶上 57 位相近，端点的 NH_2与 31 位接近。

可以预料，量子药理学在离体酶抑制剂的研究中将发挥越来越大的作用。

8.3.5 受体亚型

受体是大分子的蛋白质化合物，并非刚性结构，由于蛋白的肽链很长，容易缠卷，故而有可能同时存在几个较低能量的构象，并且这些构象之间互变的能障又很低，蛋白就容易变形，可转变为几种不同的构象，这就意味着这个受体可以以几种不同的亚型同时存在。胆碱受体分为毒蕈碱样(M 样)与烟碱样(N 样)受体，去甲肾上腺素受体区分为 α 与 β 受体，又可再分为 α_1、α_2，β_1、β_2 受体。多巴胺受体可分为 DA_1 和 DA_2 受体。组胺受体区分为 H_1 与 H_2 受体 阿片受体

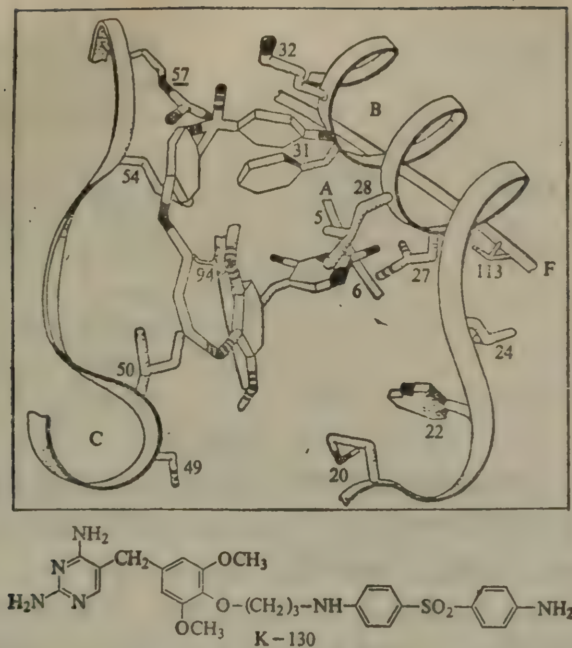
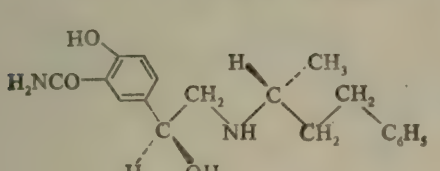


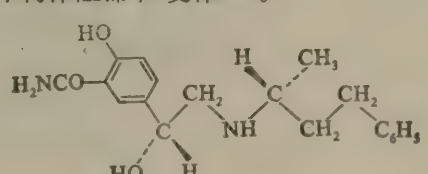
图 8-53 K-130与*E. coli* DHFR复合物示意图

至少可分为四种亚型(μ、κ、δ、σ受体亚型)。当不同结构的药物与这一受体不同构象亦即不同的受体亚型作用时,就涉及到一个分配问题。与受体产生的复合物能量愈低,就愈易紧密结合。正因如此,受体亚型的研究,也是量子药理学中构象分析和受体探索的重要应用之一[44]。

经研究肾上腺素能神经受体有两种亚型,即可分为α与β受体。其递质去甲肾上腺素由于碳链旋转,可使分子以延伸型构象与α受体结合,又可以折叠型构象与β受体作用。氯压定(6)分子也有折叠型和延伸型两种构象,表现为既有α受体激动作用,也有β受体激动作用。Labetalol(拉贝洛尔)对α与β受体都有拮抗作用。其分子带有二个手性中心,可有四个立体异构体,其中SR(1)异构体阻滞α受体,而RR(2)异构体阻滞β受体[61]。



SR- labetalol (1)



RR- labetalol (2)

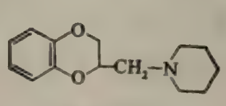
肾上腺素α与β受体又可分为α_1、α_2及β_1、β_2受体。α_1是指突触后效应受体,α_2受体本指调节递质释放的突触前受体,后来也发现突触后也有α_2受体。各种不同结构的激动剂或拮抗剂与这二受体的亲和力也发生变异[62]。图8-54指出若干α受体的激动剂与拮抗剂对α_1、α_2受体亲和力的变化。

激动剂	α_2受体	拮抗剂
氯压定 (clonidine)	↑	萝芙素 (rauvlscine)
		育亨宾碱 (yohimbine)
guanfasine		二氢麦角胺 (dihydroergotamine)
α-甲基去甲肾上腺素	亲	麦角毒碱 (ergometrine)
	和	苄唑啉 (benzazoline)
肾上腺素	力	哌氧环烷 (piperoxane)
去甲肾上腺素		苄胺唑啉 (phentolamine)
		苯氧苄胺 (phenoxybenzamine)
	↓	柯楠次碱 (corynanthine)
甲氧烷 (methoxamine)	α_1受体	哌唑嗪 (prazosin)

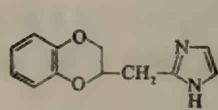
图 8-54 α_1、α_2受体亲和力变化

图 8-55 显示了α_1、α_2受体的假设受体制图。图中 A 为延伸型去甲肾上腺素分子构象示意图;B 为折叠型构象;C 是依据阿扑育亨宾碱(apoyohimbine)分子构象设想的α_1和α_2受体作用部位图。其中 1 为芳基,2 为氨基,3 为羧甲基,4 是 B 环上的氮原子。C 图表示产生α_1受体作用的必要部分是 2、3。而 1、2、4 则是α_2受体的必要部分。D 即为阿扑育亨宾碱分子,可清晰地看出其α_1和α_2受体作用部位。E 为 WB 4101 分子,F为Labetalol 的 SR 异构体。E,F 两图的分子表明它们既具有α_1受体作用,同时也具有α_2受体作用。G 为哌唑嗪分子,从它的构象分析,不具备α_2受体所需结构必要部分 1、2、4,而具备α_1受体必要部分3的结构,因而它仅对α_1受体产生作用。H 为延伸型构象的去甲肾上腺素,它具备α_1受体的必要结构,而 I 则显示了具备α_2受体必要结构的折叠型去甲肾上腺素。

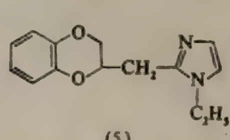
依据上述亲和力的变迁以及假设的α_1和α_2受体制图,可以有意识地改变药物的相应结构,设计更富有特异性的新药。将哌氧环烷(3)改变成 RS 21361 (4),则对α_2受体的亲和性就增加,氮上再引入乙基取代(5),对α_2受体的拮抗作用比育亨宾碱强 18 倍。



哌氧环烷(3)



RS21361(4)



(5)

氯压定(6)是α_2受体激动剂,用作降压药,但它对α_1受体也有作用,从而收缩血管。适当改变其结构,可进一步增加特异性。其分子环外的碳氮键(N^1—C^7)较短,为 0.1293 nm[63],应属双键,环内的碳氮键(N^2—C^7,N^3—C^7)的键长在 0.1325 nm,介于单键与双键之间。当质子化成盐时,由于共轭的原因,正电荷分散在各个氮原子及其所带的氢原子上。用 CND O/2 法计算,每个氮原子分有 0.17 正电荷。而类似化合物中 Lofexidine (8)与环相连的是碳而不是氮原子,

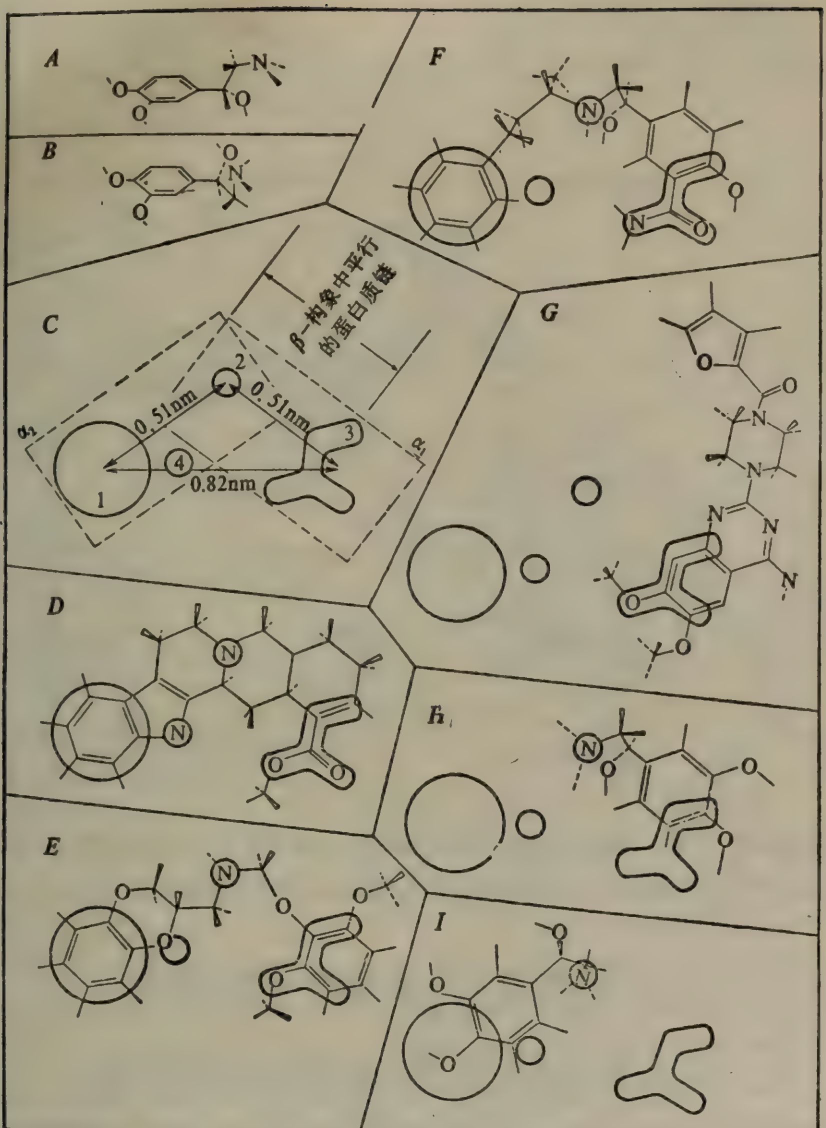
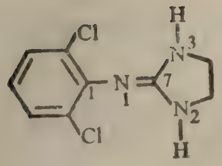
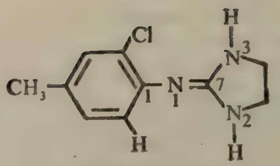


图 8-55　α_1、α_2 受体假设的受体制图

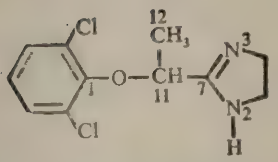
质子化后经 CNDO/2 计算每个氮原子分携 0.2 正电荷，这个化合物的作用与氯压定(6)相近。Xylazine(9)的环内硫原子不带氢原子，IPRO-4(10) 的环打开了，不再是共轭，N^2 原子在质子化后携带有 0.4 正电荷，就对 α_2 受体结合不利，因此这二个化合物作用都较弱。从构象考虑，氯压定环上二个邻位都带有取代氯，由于空间位阻使苯环与氮环不能处在同一平面，两个平面的交角(即两面角)经量子化学 PCILO 法计算为 75°，与 X 衍射结果吻合。Tolnidine 的邻位只有一个取代氯，交角就较小($\phi=55°$)，而 Lofexidine 的二个取代氯加上侧链上的取



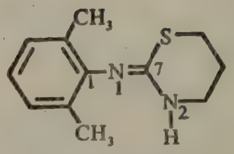
clonidine(6)



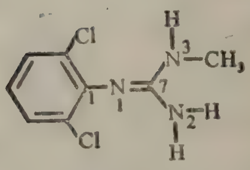
tolonidine(7)



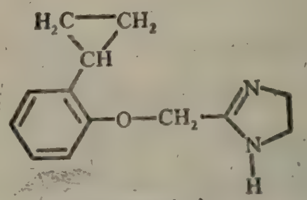
lofexidine (8)



xylazine (9)



IPRO-4 (10)



cirazoline(11)

代甲基，就又增加了交角。若结构进一步改变成Cirazoline(11)，则变成α_1受体激动剂，对α_2受体则有拮抗作用，从而使(11)在临床上的用途转变化为缓解鼻充血。

8.4 药物分子的静电势和电荷密度

8.4.1 量子药理学中静电势和电荷密度的计算

如8.3中所述，按照药物分子内电荷密度分布，可以反推受体各相应作用点的性质。电荷密度较高的部位(即带有较高负电荷)，可作用于受体的正电中心；而电荷密度较低的部位，则可作用于受体的负电中心。静电势可描绘分子周围的电荷作用，负电区域有利于受体正电部位结合，反之亦然。

在用分子轨道方法计算药物分子电子结构时，一般若药物分子系刚性结构，则可用该分子X衍射测得的晶体结构，作为它的稳定构象或药效构象。将X衍射测得的空间各原子的分数坐标，经变换后即可获得量子化学计算的直角坐标。或者依据结晶报告中公布的键长、键角、二面角数据，依据右手法则构筑分子内坐标也可进行量子化学计算。有的结晶报告不公布氢原子坐标，亦可借助“加氢程序”方便地给出。图8-56显示了我国自行开发研制的新型化学结构的抗疟药

——青蒿素(Artemisinine)的透视图。它是一刚性结构的药物。可直接用其X衍射结晶数据，将青蒿素的结晶分数坐标值按所示的空间群 $p2_12_12_1$ 的晶胞中各分子的对应值。一一计算获各原子的直角坐标。即可进行分子轨道计算，图 8-57 显示了用 EHMO 法计算获得的青蒿素各原子的电荷密度值和原子间的键级[64]。

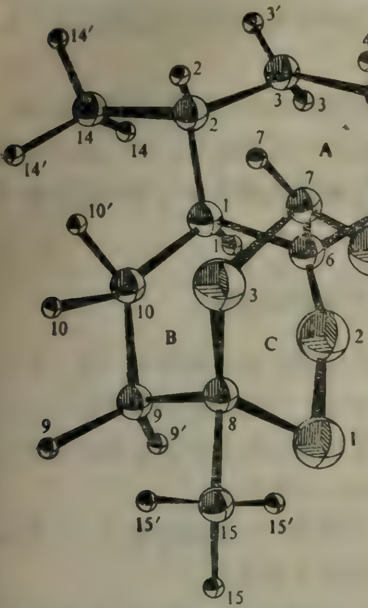


图 8-56 青蒿素分子透视图

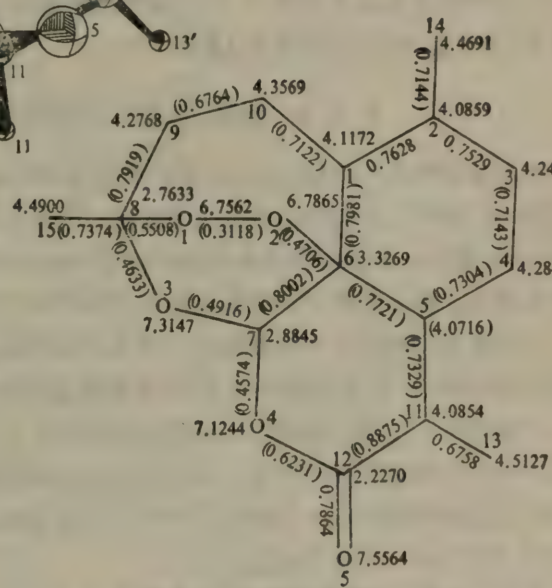


图 8-57 青蒿素分子的电荷分布

有时我们处理的对象是尚未获得X衍射结晶数据的化合物。但又需先进行量子化学计算。在这种情况下，我们可采取标准键长、键角、二面角的办法，或者主要结构部分查找相应的X衍射结晶报告，再拼以相连原子的办法来近似处理。这在进行半经验分子轨道法计算时，尤其是计算同源化合物系列分子时，尚不致引起很大误差。

对于柔性药物分子，由于有多个扭角可以旋转，一般除参照X衍射结晶或NMR数据外，需先进行构象分析，将优化后所获能量最低的构象看成是稳定构象，在当今对多数药物受体缺乏了解的情况下，就把这一能量最低的构象称为药效构象或称药效基团。关于构象计算 8.2 中已有详述。事实上，量子化学方法中绝大多数方法在计算能量后即可同时获得各原子的电荷密度，甚至进而获得前沿轨道指数，如前沿轨道电荷密度、超离域度等。因此所获稳定构象后即同时产生所需电荷密度等指数。近年来，由于量子化学计算程序的日趋完善，加上各种计算机的功能的扩大，许多先进的量子化学程序都自动包含有能量优化的部分。如从 QCPE 获得的 Gaussian 80 和 Gaussian 82 从头计算程序除可计算常用的各种展开基组外，都包含有能量梯度优化和 CI(组态

相互作用)。半经验的 MNDO 法也采用了能量梯度优化。不过由于药物分子较大,目前还只能限于MNDO计算的优化上,而从头计算优化还不普遍,往往限于较小分子或片断模型分子,这时往往事先采用PCILO或分子力学方法优化构象,再用稳定构象进行整个分子的从头计算法计算。常常为节省机时 (computer time) 采用整个分子取不同基组展开的技术,如影响活性的主要部位的原子采用高基组(如 3-21G或 4-31G),次要部位或同源化合物中结构相同的部位采用一般基组(如 STO-3G甚至 STO-2G)[48]。

分子静电势的计算,最常用的是 ab initio 法和 CNDO/2 法获得的波函数来计算静电势。(详见第二章)值得一提的是,分子片断模型的处理常为分子静电势计算所采用,ab initio 法如此,连 CNDO/2 法也不例外[65-67]。为了节省机时,从头算中越来越多的静电势计算采取价电子从头算(赝势法)的方法(LP-3G基组)[68,69]。

8.4.2 组胺及其衍生物的电荷密度和分子静电势

组胺是动物体内广泛分布、大量存在的重要内源性化学介质,它由组氨酸脱羧后产生。它作为一种神经递质能引起体内多种生理效应。众所周知,动物体内至少存在两种组胺受体——H_1受体和 H_2 受体。组胺对这两种受体都有强烈的作用,而组胺的某些类似化合物却显现出明显的受体选择性。量子药理学对组胺构象及受体作用已进行了大量深入的研究。

我们知道,组胺、2-甲基组胺和 4-甲基组胺是乙胺基咪唑类化合物。2-甲基组胺对 H_1 受体有较强的作用,对 H_2受体作用甚弱;4-甲基组胺的受体作用刚好相反。见表 8-10。

实验表明[70]:在水溶液中,这些化合物形成了一个由负离子、中性分子、一价和二价阳离子的多种异构体组成的平衡体系。(见图 8-58),在人体生理pH值(pH=7.4)下,浓度最大,对组胺的活性起主要作用的是单价阳离子的N^{τ}—H互变体。在用量子化学计算讨论组胺与受体的作用时,往往采用 N^{τ}—H 型互变体。

为探讨组胺、2-甲基组胺、4-甲基组胺对 H_1、H_2 受体的作用,陈凯先等[71]用从头计算法,采用STO-2G基组对上述三个化合物进行了计算,获得了分子轨道能量及电荷密度,以及根据从头计算所获的波函数进行分子静电势的计算。图 8-59 表示三个分子各原子的净电荷值。

通常,原子上的净电荷是通过平分两原子间电子重叠集居数获得的,它不考虑各原子电负性的差别,因而不可避免地存在着缺陷。但是分子静电势计算及据此绘制的分子静电势图则可提供分子静电性质的准确、客观和直观的图象。由于在药物与受体作用的许多场合,静电力是一种重要的因素,因此静电势计算受到了量子药理学家的重视。图 8-60—图 8-62 显示了组胺在三

表 8-10 E_{HOMO}和E_{LUMO}(kcal/mol)和H_1、H_2受体活性比较

	组　胺	2-甲基组胺	4-甲基组胺		组　胺	4-甲基组胺	2-甲基组胺
E_{HOMO}	−224.837	−216.742	−216.052	E_{LUMO}	146.461	149.159	150.163
能级次序	——————→			能级次序	——————→		
H_1受体活性	←——————			H_2受体活性	←——————		

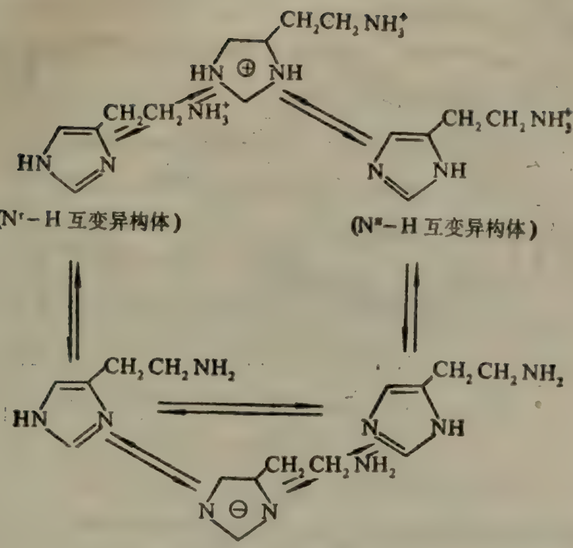


图 8-58 溶液中组胺的平衡体系

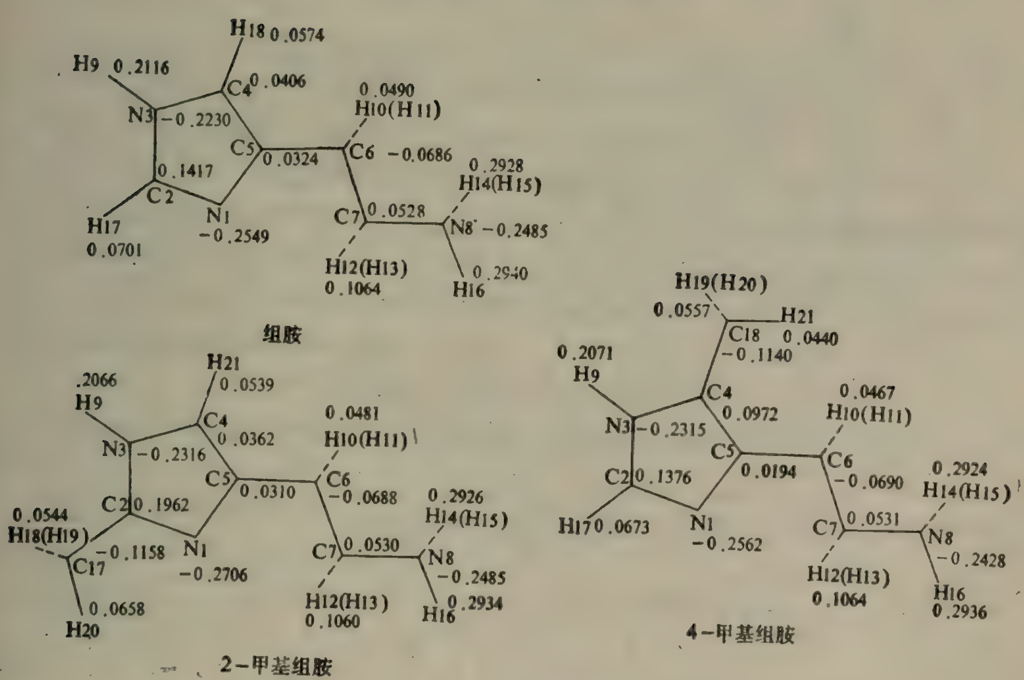


图 8-59 原子上的净电荷

个不同平面 $x=0$(即咪唑环所在平面);$x=0.75$ a.u.和 $x=1.5$ a.u.上的静电势图形,从中可以了解组胺分子静电势的立体图象。图 8-63、图 8-64 分别是 2-甲基组胺和 4-甲基组胺在咪唑环平面内的分子静电势图。

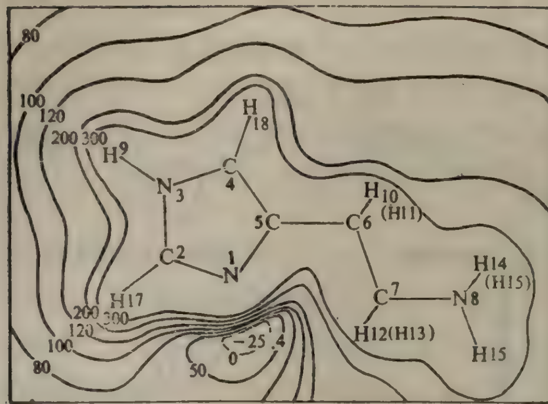


图 8-60　组胺(阳离子)静电势图(咪唑环平面*内)

* 咪唑环平面即 $y—z$ 平面

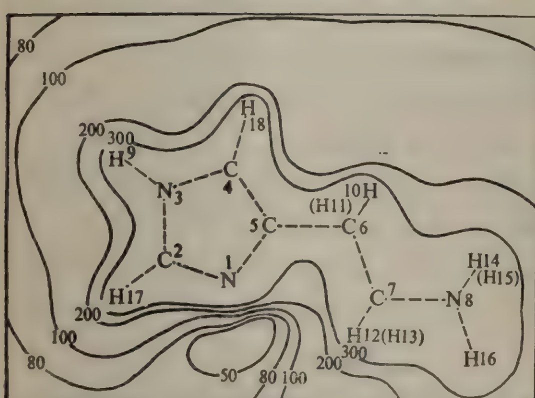


图 8-61　组胺(阳离子)静电势图($x=0.75$ a.u.平面)

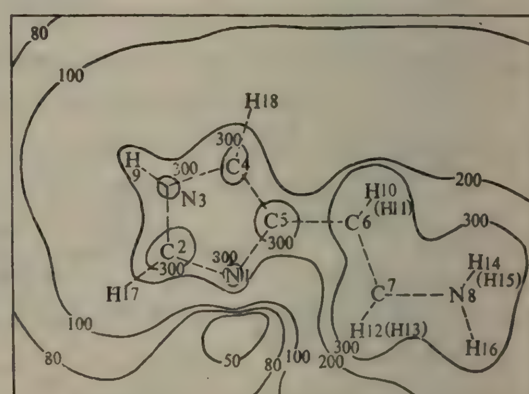


图 8-62　组胺(阳离子)静电势图($x=1.50$ a.u.平面)

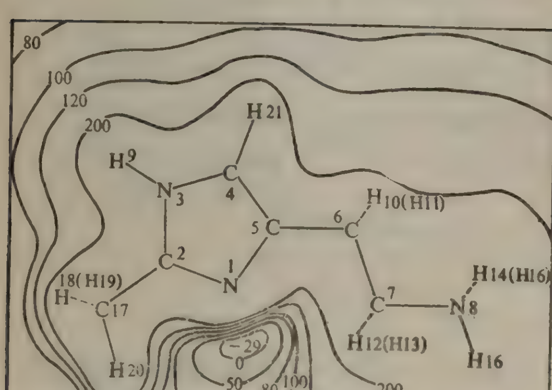


图 8-63　2-甲基组胺(阳离子)静电势图(咪唑环平面内)

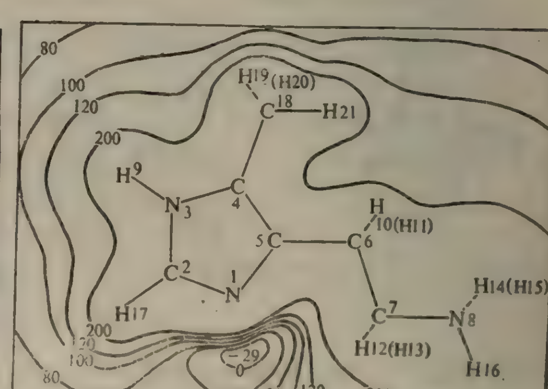


图 8-64　4-甲基组胺(阳离子)静电势图(咪唑环平面内)

通过分析计算结果可知，首先在组胺及其衍生物的生物活性与各自的前线分子轨道能级之间存在着一种对应关系。表 8-10表明，随着最高占有轨道能级(E_{HOMO})降低，H_1 受体活性增强，随着最低空轨道能级(E_{LUMO})降低，H_2 受体活性增强。

从占据轨道和空轨道的波函数可看出，最高占有轨道和最低空轨道，主要都是由咪唑环上碳原子和氮原子的垂直于环平面的 p_z 轨道所组成。所不同的是，最高占有轨道是一成键的分子轨道，而最低空轨道则是反键轨道。从表 8-10可知，以 E_{HOMO}为标志的成键轨道上的电子稳定性，影响着该化合物的 H_1受体；而以 E_{LUMO} 为标志的反键轨道获得电子的能力则与该化合物的 H_2 受体活性相关联。这表明，在组胺产生生物活性的过程中，受体与咪唑环通过电子迁移过程和静电相互作用的结合，可能是重要因素。咪唑环2、3、4位如引入甲基，由于甲基的空间位阻，使受体与咪唑环的结合受到影响。这三个化合物的 H_1、H_2受体活性经过实验，确实远比组胺为低。表 8-10显示组胺与衍生物的 E_{HOMO}相差8kcal/mol 以上，E_{LUMO}相差 3 kcal/mol 以上，远比衍生物之间相差 1 kcal/mol 为大。

从分子静电势图可以发现，除了咪唑环 1 位上氮原子附近以外，分子的其余部分均被正电势的套所包围。这一图象表明在距受体较远时，组胺阳离子就整体而言，是一个带正电的实体，由此推测其受体应带一定的负电荷。它们在 Columb 力作用下彼此接近时，组胺及衍生物分子静电势的细节，就会被受体进一步感知和辨认。

咪唑环 1 位氮原子附近的负区域对于 H_2受体活性可能具有影响。当2-位引入甲基后，这一负区域被甲基的正区域所部分屏障，这将影响 H_2受体中相应的正的部分与这一负区域的契合。这可能是2-甲基组胺的 H_2受体活性远比组胺和4-甲基组胺为低的一个因素。

侧链末端为静电势最高区域，它和咪唑环 1 位氮原子附近的负区域构成了分子静电势的正、负中心，其间距离约为 0.5 nm。这一静电势特征可能与 H_2受体的作用有关。甲氰咪胍 (cimetidine)侧链虽较长，但由于分子内氢键的作用，侧链发生了折叠，并使其侧链胍基上氮原子与咪唑环 1 位氮原子的距离仍符合上述静电势特征，故可与 H_2受体结合成为 H_2受体拮抗剂，这便是上述推测的印证之一。

2-甲基组胺与4-甲基组胺的差别，仅在于甲基取代的位置不同，但两者 H_1 受体活性相差达85倍。仔细分析两种分子构象的差别以及由此产生的静电势差异，也应能反映其 H_1受体活性差别的原因。

根据 ab initio 法计算获得的能量、电荷密度和静电势，我们可以进一步推测受体的大致图象。组胺受体的一些可能的特征是：

1） 就整体而言，它应是显示较强的负电性。

2） 应具有和咪唑环配合的疏水区。

3） 几何形状和静电势性质应和组胺的 H_1、H_2受体活性构象充分互补。

8.4.3 β 受体药物的分子静电势研究

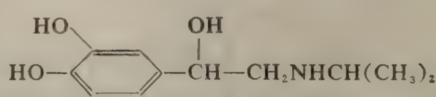
在肾上腺能神经的研究中，已发现有两种受体，即 α 与 β 受体。其中 β 受体的兴奋会产生心搏增快、心脏收缩增强，以及支气管舒张等生理效应。拮抗 β 受体便可产生相反的效应，可促使

心搏减缓、心脏收缩减弱及支气管收缩等。

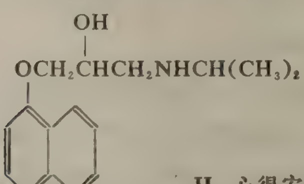
异丙肾上腺素(I)是标准的β受体兴奋药物，它的结构特征是苯乙胺型化合物，其他一些具有苯乙胺型结构的化合物也具有β受体兴奋作用。心得安(II)是常用的β受体拮抗剂，它的结构特征是苯氧丙胺类化合物，许多类似结构的化合物也有拮抗作用。如甲苯心安(Doberol，III)便是一例。

随着研究的深入，发现有的具有苯乙胺型结构的化合物反而成为β受体拮抗剂，如硝苯心定(nifenolol IV)。而另一苯氧丙胺型化合物塔唑醇(tazolol，V)反而有β受体兴奋剂的作用，用于治疗心力衰竭。

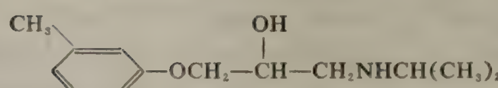
为深入探讨β受体兴奋剂与拮抗剂的结构活性关系，Petrongolo 等对有关化合物进行了 CNDO/2 和 ab initio 计算，并依据 ab initio (STO-3 G) 所获的波函数进行了分子静电势研究[72,73]。



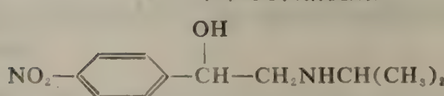
I. 异丙肾上腺素(兴奋剂)



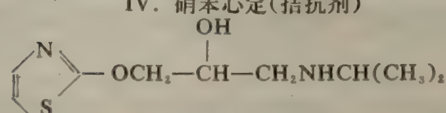
II. 心得安(拮抗剂)



III. 甲苯心安(拮抗剂)



IV. 硝苯心定(拮抗剂)



V. 塔唑醇(兴奋剂)

在静电势计算中，为简化计算，均将其侧链简化，并以模型分子 IA 来代替。图 8-65 系 IA 分子在苯环平面上的分子静电势图，图中静电势值系 10^{-3}a. u. 单位。

为了更好地考察苯环周围的静电势变化，而后的计算均在距离为 0.17 nm 且平行于苯环的平面(即 $z=0.17$ nm)上进行，并在 C 7 位置上引进 O 4 原子。取 $z=0.17$ nm 的平面是由于这个平面上的静电势图可较明显地体现 π 电子的贡献。

图8-66—图8-69分别为分子 I、III、IV、V 的静电势图。

从四个分子静电势图可知，异丙基肾上腺素三个氧原子周围是较强的负电势区。苯环平面

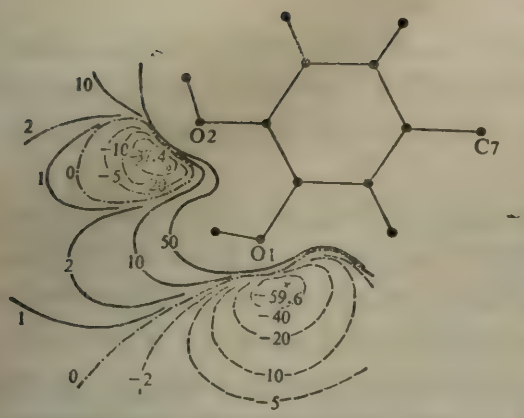


图 8-65　IA模型分子苯环平面上的分子静电势图

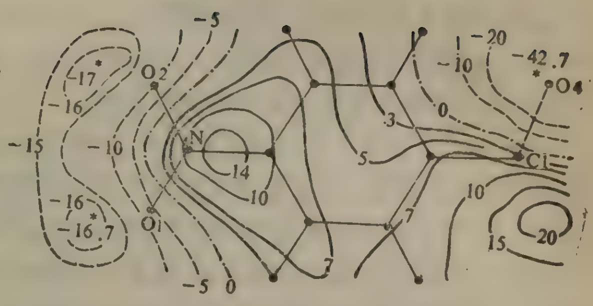


图 8-67　分子IV的分子静电势图

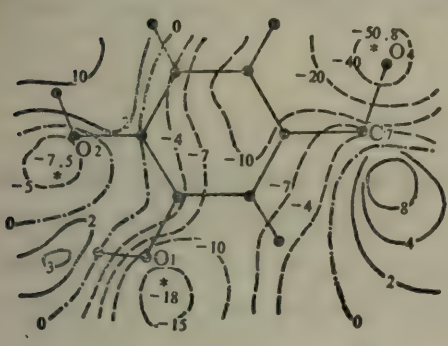


图 8-66　分子I的模型分子，在 $z=0.17$nm 平面上的分子静电势图(10^{-3}a.u.)(下同)

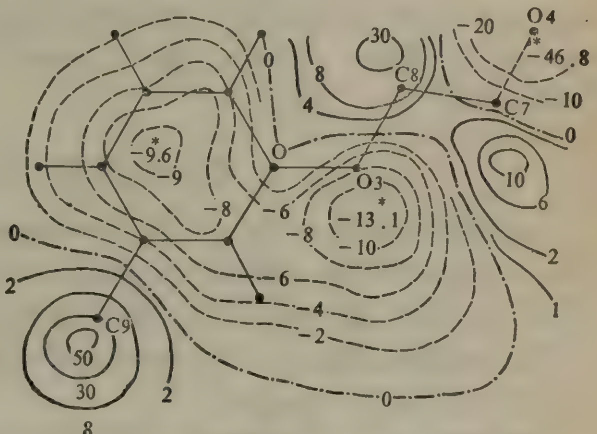


图 8-68　分子III的分子静电势图

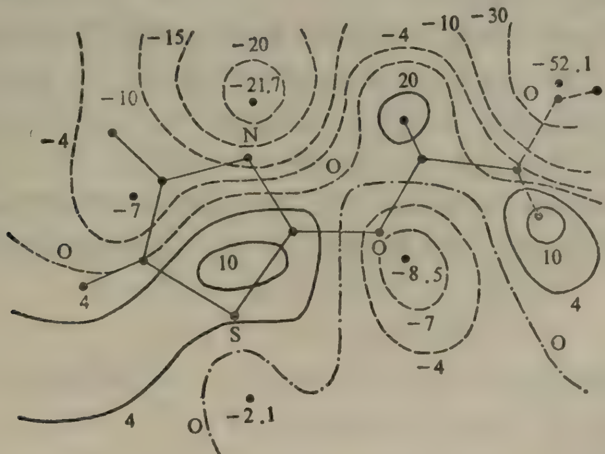


图 8-69　分子V的分子静电势图

也为一负电势区所包围(见图8-66),因而显示苯环附近为亲电区域(electrophilic attack),这对苯环的亲核进攻(或反应)是有利的。在硝苯心定分子中,硝基的两个氧原子周围是一强的负电势区,由于硝基的作用,致使苯环周围成为正电势区(见图8-68),变作亲核区域,对其亲核反应不利,而易于发生亲电反应(或进攻),这是从静电势计算分析所得异丙肾上腺素是β受体兴奋剂,而硝苯心定则是β受体拮抗剂的主要原因。

图8-67显示了甲苯心安的静电势,可见其环上间位甲基附近存在强的正电势区,而侧链上氧原子旁的次甲基分开了两个负电势区,次甲基本身也是正电势区。这就与硝苯心定的静电势分布类似,即它的负区域苯环相当于后者的硝基周围的电势,是一个负区域,而它的次甲基相当于后者的苯环周围,是一个正区域,这是β受体拮抗剂的共同的静电势特征。图8-69显示了塔唑醇分子的静电势,由于环内氮原子上孤对电子的作用,使其周围形成一负电势区,其位置相当于异丙肾上腺素对位羟基上氧原子周围的负电势,而侧链上氧原子附近的负电势则相当于后者的苯环,为一负电势区。由此可见,在β受体兴奋剂中有一个负电势区,为亲电区域。而在β受体拮抗剂中有着一个正电势区。这些区域在苯乙胺型化合物中存在于苯环,在苯氧丙胺型化合物中却存在于侧链的$—OCH_2—$。实际上,这就从量子药理学角度解释了定性的构效关系所不能回答的问题。

8.4.4 增敏剂的量子药理学研究

辐射增敏剂(radiosensitizer)是一些化学药物,它们能加强射线对细胞的损伤作用。肿瘤放疗中面临的最严重的问题是肿瘤的复发和射线对正常组织的损伤作用,一般认为,肿瘤的放疗所以不一定奏效,是由于肿瘤中存在对射线有较强抵抗力的乏氧细胞之故。如在放疗中同时施以乏氧细胞的辐射增敏剂,可选择性地使乏氧细胞敏化,从而提高放疗临床效果,所以增敏剂的合成与研究引起了广泛的注意。

随着研究的深入,促进有效增敏剂的设计合成和药理筛选已经提到议事日程上,随之对增敏剂的作用机制的探索也逐步开展起来。

Adams等提出了化学增敏剂电子转移模型和直接作用模型[74],把增敏现象与增敏剂的电子性质紧密地联系起来,后又提出由脉冲射解测得的单电子还原电位与增敏活性存在明显的相关性[75],进而提出亲电理论的机制,大大促进了增敏机制的研究。以后又有自由基增敏机制的产生。近年来,增敏机制的探索也成为量子药理学的一个重要内容。

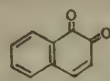
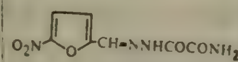
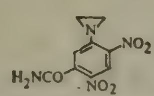
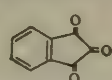
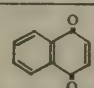
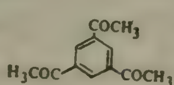
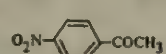
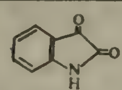
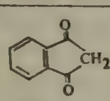
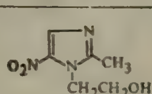
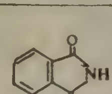
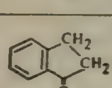
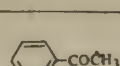
徐素娟等[76]采用CNDO/2方法对13个常用的一系列分子母体各异的乏氧细胞辐射增敏剂进行了计算。结果发现,增敏剂的生物活性与$S_{max}^{(R)}$这一量化指数之间存在良好的相关性。

$$\log 1/C_{1.6}=-0.69+1.78\log(S_{max}^{(R)}-3) \tag{8.7}$$

$$n=13;\ r=0.93;\ s=0.36$$

式中$C_{1.6}$是SER(sensitizing enhancement ratio)等于1.6时增敏剂的克分子浓度,它代表了增敏剂的生物活性浓度。$S_r^{(R)}$是亲自由基超离域度,由CNDO/2计算所获,它是分子内各位点与自由基反应难易程度的尺度,也是比较分子之间与自由基反应能力大小的指数。$S_{max}^{(R)}$系分子中最大的$S_r^{(R)}$值。表8-11是所计算的增敏剂分子,图8-70表明$\log 1/C_{1.6}$与$S_{max}^{(R)}$之间的指数

表 8-11　乏氧细胞辐射增敏剂及其超离域度和增敏活性

编号	化合物	结构式	$S_{max}^{(R)}$	增敏效果		$C_{1.0}$	$\log(1/C_{1.0})$	
				nmol·l⁻¹	SER		实验值	计算值
1	1.2-萘醌		61.3	0.006	1.6	0.006	2.22	2.45
2	硝基呋喃草酰腙		18.1	0.25	2.9	0.029	1.54	1.41
3	2.4-二硝基-吖丙啶苯酰胺		11.7	0.2	2.2	0.05	1.30	0.98
4	茚满三酮		7.9	1.0	3.0	0.13	0.89	0.54
5	1,4-萘醌		10.4	0.1	1.54	0.15	0.82	0.86
6	1,3,5-三乙酰苯		5.8	0.3	1.8	0.2	0.70	0.11
7	对硝基苯乙酮		8.2	0.4	1.7	0.3	0.52	0.59
8	吲哚醌		6.0	5.0	2.4	0.7*	0.15	0.16
9	1,3-茚满二酮		4.5	0.5 1.0 3.0	1.0 1.1 2.1	2.0*	−0.30	−0.37
10	灭滴灵		5.7	10.0 0.25	1.9 1.25	4.0	−0.60	0.08
11	酞酰亚胺		5.2	1.0	1.0	4.0**	−0.60	−0.08
12	茚满酮		4.0	2.0	1.1	4.0**	−0.60	−0.69
13	乙酰苯		4.1	20.0	1.0	4.0**	−0.60	−0.61

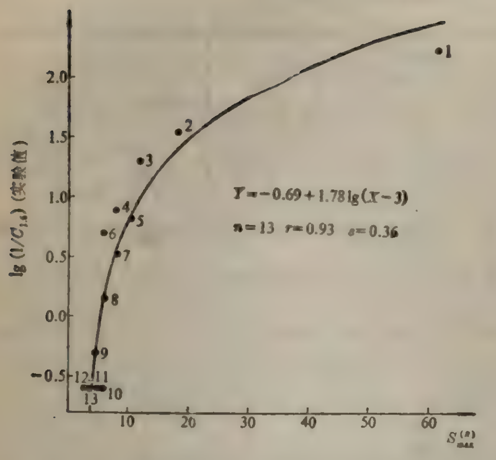


图 8-70 lg1/$C_{1.6}$与$S_{max}^{(R)}$关系图

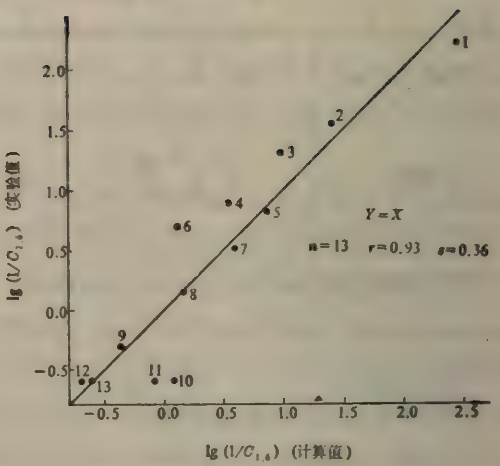


图 8-71

型相关关系.图 8-71 显示实验所获的 log 1/$C_{1.6}$与根据式(8.7) 所计算的 log 1/$C_{1.6}$值的比较，显良好的线性关系。

这一计算结果从量子药理的角度上说明自由基反应在增敏过程中的重要性,在一定程度上支持了乏氧细胞增敏剂通过与损伤的靶分子自由基直接发生反应而引起增敏作用的机制。

计算结果还表明,增敏活性与增敏剂分子的最低空轨道能量 E_{LUMO} 存在线性关系,E_{LUMO}愈低,增敏活性越高,(见图 8-72)。增敏活性还与$S_{max}^{(N)}$值成线性关系。E_{LUMO}值代表了化合物亲电能力,其值愈低表明亲电能力愈强。$S_r^{(N)}$是表示亲核反应能力大小,若讨论该分子与受体作用,则表示其亲电进攻的能力。$S_{max}^{(N)}$与增敏活性呈正相关,其值愈大,活性愈强。这一分析表明，量子药理学计算进一步支持了亲电理论的直接作用模型。

事实上,Willson等经实验证明,乏氧细胞增敏剂的增敏作用中至少可包含四种增敏机制,其

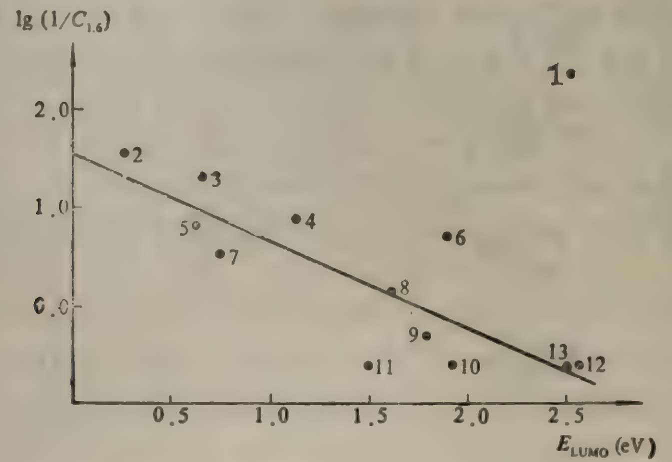


图8-72 log1/$C_{1.6}$与 E_{LUMO} 之间的相关性

中就包括亲电的直接作用模型和通过与损伤分子自由基反应的机制[77]。徐素娟等在对量子化学 CNDO/2 计算所获的信息(即$S^{(N)}_{max}$和$S^{(R)}_{max}$等指数)进行了对照分析，按照增敏活性强弱范围的判据标准，从理论上预计汉防己甲基-双苄基异喹啉生物碱(bisbenzy lisoquinoline, alkaloid)及其类似物，认为它们不具有乏氧细胞的增敏活性[78]，经细胞试验及临床观察表明理论判据是正确的[79]，从而避免了反复试验和大量的对照观察。

通过对 11 个具有平面杂环的分子进行分子静电势计算，(所用的方法是经 CNDO/2 获得的波函数进行静电势计算)，也获得了对增敏机制探索的一些有益结果[80]。

图 8-73 至图 8-76 显示了九元环共轭杂环衍生物的分子静电势图，我们可以看出茚满三酮(triketoidane)的静电势图中有一处于同一侧的、连续的、深而宽的复区域。吲哚醌(isatin)也在同一侧有连续的负区域。这两个分子都有较好的增敏活性。而 1,3-茚满二酮(1,3-indanedione)和酞酰亚胺(phthalimide)两分子的负区域被正电势区隔开，处于不连续的两侧位置，而它们的增敏活性较差。

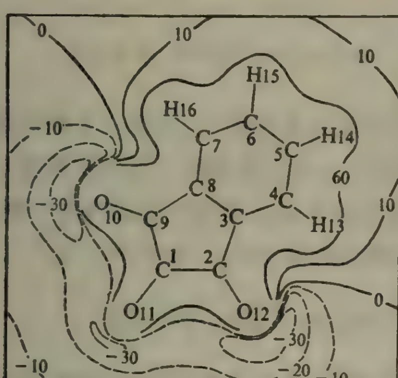


图 8-73　茚满三酮分子静电势图

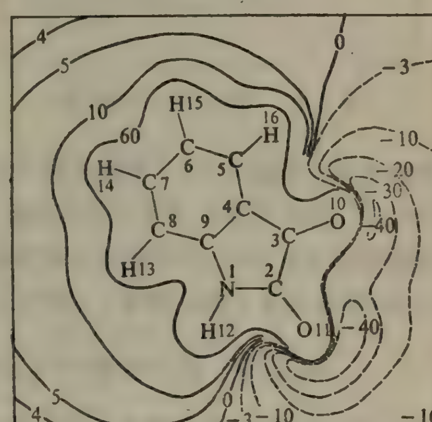


图 8-74　吲哚醌分子静电势图

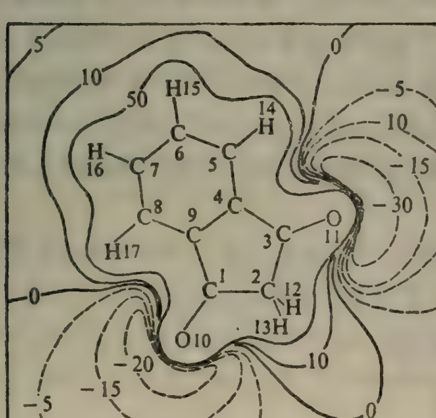


图 8-75　1,3-茚满二酮分子静电势图

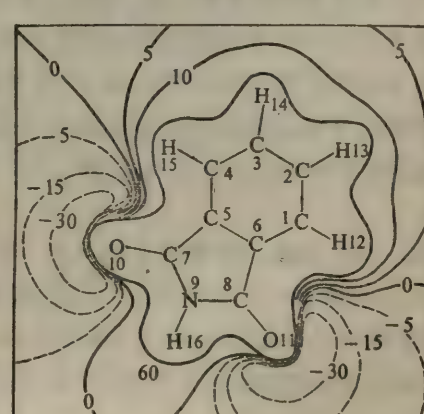


图 8-76　酞酰亚胺分子静电势图

图8-77，图8-78显示了十元环的萘衍生物，两个分子只是由于一个羰基的位置从原来的1,4-萘醌(1,4-naphthaquinone)的4位移至2位，成1,2-萘醌(1,2-naphthaquinone)，使原来不连续处于对侧的负电势变成连续的同一侧的负电势区域，致使增敏活性大大增加 (log $1/C_{1.6}$由0.82增至2.22)。

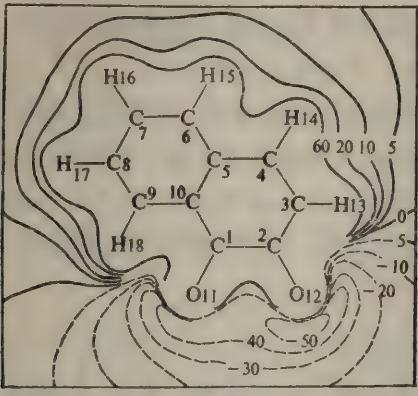


图 8-77 1,2-萘醌分子静电势图

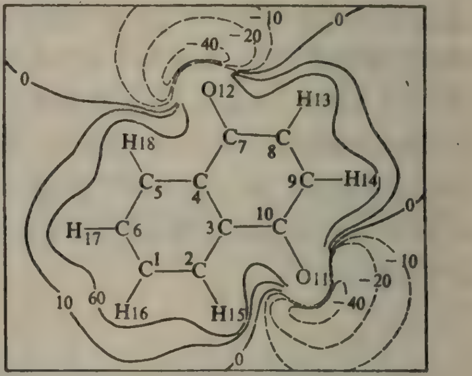


图 8-78 1,4-萘醌分子静电势图

上述静电势计算提示人们，增敏活性不仅与分子中的活性基团有关，而且与这些基团的位置有一定的关系。活性基团在分子中的排列使分子平面形成一个同侧连续，深而宽的负区域的增敏剂具有较高的增敏活性，这无疑是徐素娟等提出的对增敏机制及增敏剂筛选的又一理论判据。随后Inayma等也报导了他们使用CNDO/2分子静电势的研究，获得了相同的结论[81]。

8.4.5 苯甲酰胺型D_2受体拮抗剂的分子静电势计算

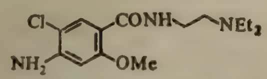
自从Carlsson于1958年报道纹状体内多巴胺(DA)的含量占全脑的70%以来，引起了神经药理工作者的很大兴趣。60年代证明巴金森氏症(Parkinson's disease,PD)的病因是由于黑质致密区DA神经元的变性所引起，能用1-多巴(dopa)替代治疗，这些结果是当时神经递质药理学划时代的成果。1975年应用放射受体结合法证实了配基能直接与DA受体的特定受点相结合，进一步深入的研究表明脑内有不同的DA受体亚型，目前比较倾向于D_1和D_2两种亚型的分类。

苯甲酰胺型化合物是较近发展的D_2拮抗剂，如舒必利(sulpiride)、胃复安(metoclopramide)、硫必利(tiapride)、clebopride等。它们具有抗精神病、抗呕吐、抗运动障碍和安定等神经药理作用。

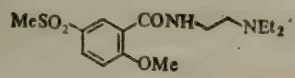
对这一类D_2拮抗剂的研究自然引起了量子药理学工作者的兴趣。Waterbeemd等运用量子化学PCILO法辅以分子力学MM 1计算，研究了一系列苯甲酰胺型化合物的优势构象，其主要特征为，酰胺侧链的邻位甲氧基，可与酰胺的氢产生氢键。其中舒必利的氨基氮原子与苯环中心相距0.62 nm，胃复安的氮原子与苯环中心相距0.595 nm[82]，与延伸型多巴胺的相应距离(0.60 nm)吻合。

为进一步研究其与受体作用的情形，Waterbeemd从苯甲酰胺型化合物的优势构象出发，进

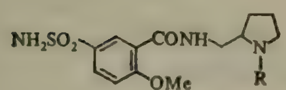
行分子静电势研究，他们整理了有明确药理作用的苯甲酰胺型化合物共 11 个，其结构式与命名见图 8-79。



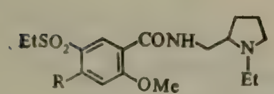
metoclopramide



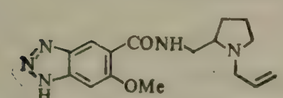
tiapride



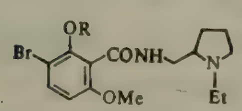
R＝Et sulpiride
R＝CH_2-C_6H_4-P-F flubepride



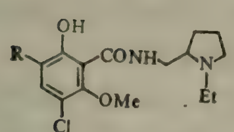
R＝H sultopride
R＝NH_2 amisulpride



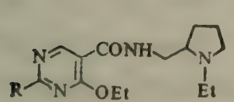
alizapride



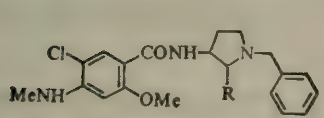
R＝H FLA 797
R＝Me remoxipride



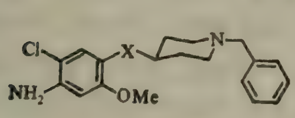
R＝Et eticlopride
R＝Cl raclopride



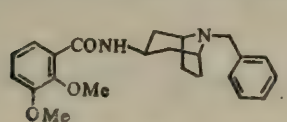
R＝Me
R＝NH_2



R＝H YM 08050
R＝Me YM 09151-2



X＝CONH clebopride
X＝NHCO BRL 20596



tropapride

图 8-79 苯甲酰胺型化合物结构式

由于分子较大，他们将侧链裁减，综合成 8 个模型化合物（图 8-80），这 8 个化合物都可表示

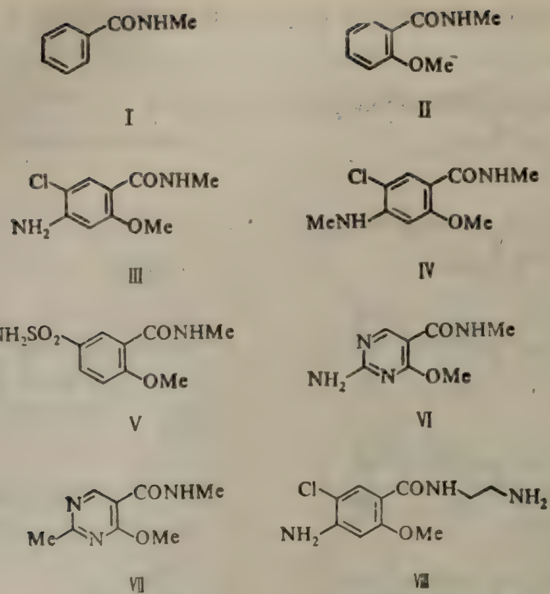


图 8-80 模拟分子结构式图

前述 11 个化合物芳香环上甲氧基与邻位的酰胺基的作用。作者选用从头计算法，采用STO-3 G极小基组，使用MONTER-GAUSS程序在CDC CYBER 170/720 计算机上进行。用从头算获得的波函数进行分子静电势计算，使用 QCPE 的 VSS 程序(QCPE 245)。

考虑到要研究 π 电子分布对整个静电势的影响，因此，选取平行于芳香环 0.175 nm 的平面上计算分子静电势，图 8-81 显示了化合物I—VII分子的静电势图。

从图 8-81 可见，除化合物 I 因侧链没有甲氧基取代外，该类药物的分子静电势基本由两部分组成。由甲氧基氧原子与侧链酰胺基上氮原子形成的六员环为正电势区域，羰基氧及其同侧的取代基(如—NH_2、Cl) 形成负电势区域。两个区之间有一条基本上穿越芳香环的分界线。由于化合物VIII的侧链较长，又最接近于胃复安的结构，因此作者进一步选取平行于苯环四种不同的距离的平面，(分别为 0.175 nm，0.20 nm，0.25 nm，0.30 nm)，计算分子静电势。图8-82 (370页)显示了4 种不同高度的静电势图。

从图 8-82 可见，在苯环上方 0.30 nm的平行面上，正负静电势区域的分界线比较平坦，随着这个平行面向苯环逐渐接近，分界线就愈来愈弯曲。这就表明，当药物分子与受体相互作用时，首先是按正负电势区域要求排列及取向，然后二者进一步靠近，受体点进行识别，结合，形成药物受体复合物，产生药理作用。这类化合物的静电势区域也与多巴胺的静电势区域相吻合。

8.5 电荷转移

8.5.1 电荷转移中的量子药理学研究

电荷转移是生物体系的重要作用方式之一。在电荷转移过程中，一个给予体分子(D) 转移

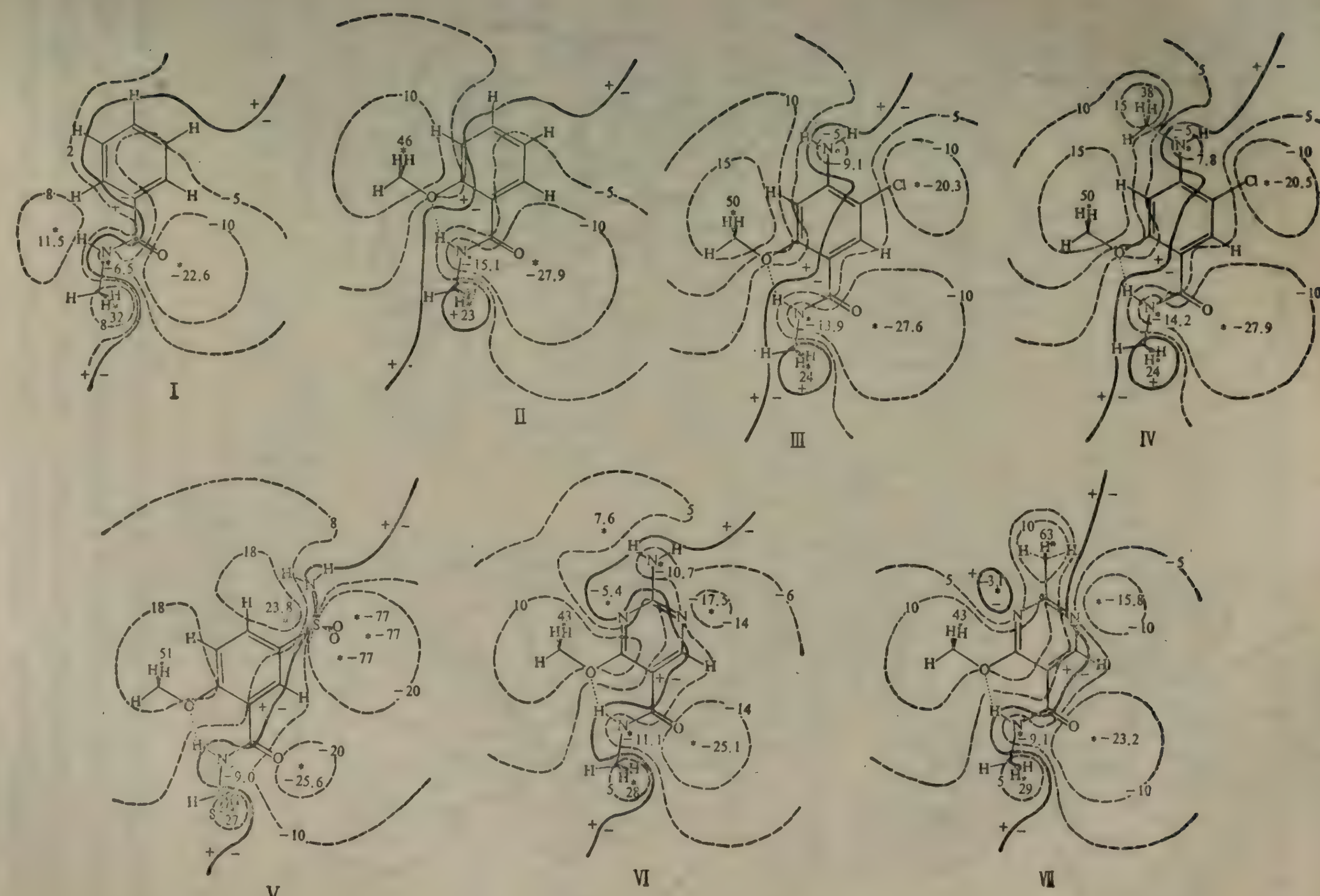


图 8-81　模拟分子 I—VII 分子静电势图(芳环上方 0.175 nm 平行面 kcal/mol)

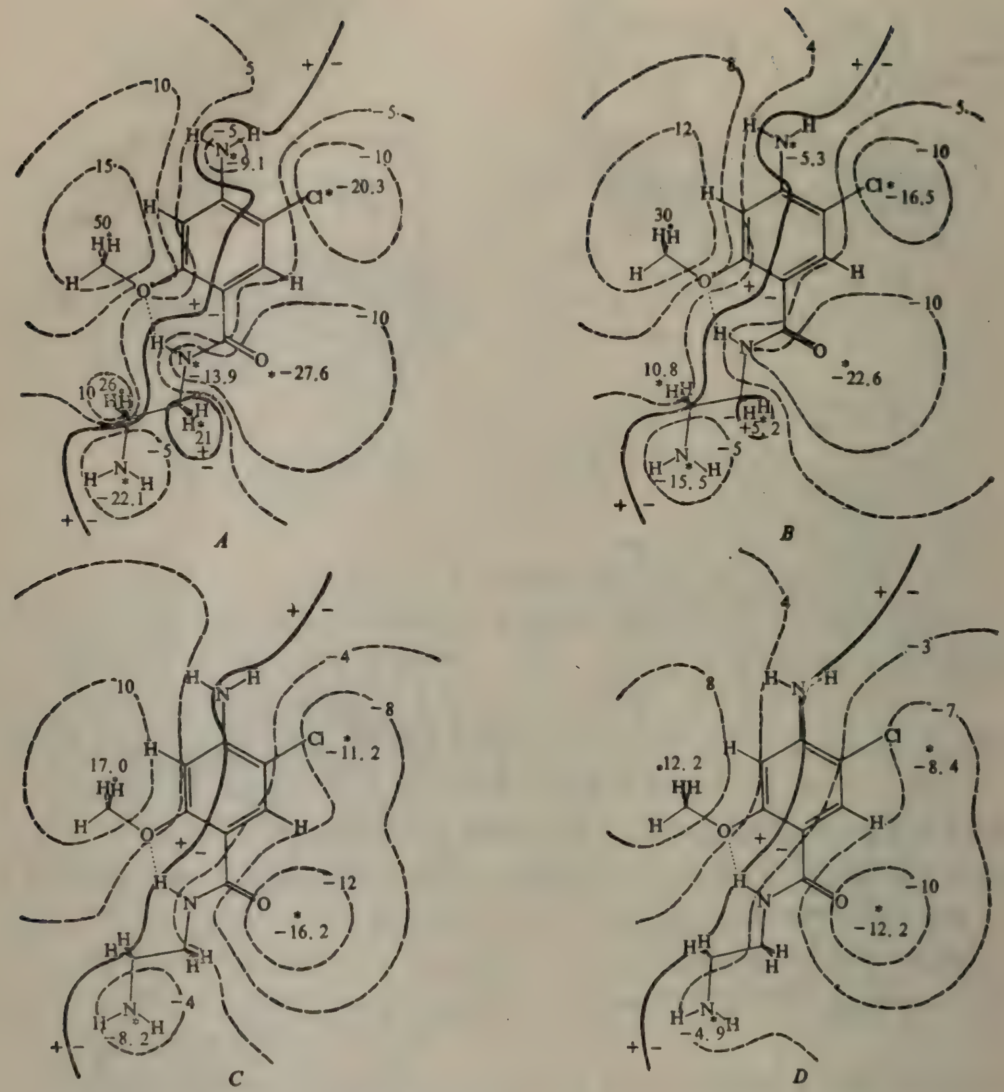


图 8-82　模拟分子 VIII 在芳环上方平行面 *A*．0.175 nm；*B*．0.20 nm；*C*．0.25 nm；*D*．0.30 nm 上的分子静电势(kcal/mol)图

一个电子至接受体分子(A)，然后相互复合：

$$D + A \longrightarrow D^{+} + A^{-} \longrightarrow D^{+}A^{-}$$

这种复合体称为“电荷转移复合物”，这种结合力较强，足以产生吸收光谱。由于大多数药物与受体的结合是可逆的，亦即不能通过产生共价键的途径，而偶极矩间结合或范德华引力的作用都较弱，故通过这些作用还不能使药物与受体结合。但通过电荷转移，以形成电荷转移复合物的方式使药物与受体结合，可能是一重要途径，因而有关电荷转移的研究也成为量子药理学中相当活跃的课题。

我们知道，只有电子易于流动的体系，才能产生电荷转移。给予体的电子来源，可来自非键的孤对电子，如胺、醚、醇、碘化物的孤对电子，称 n 给予体，也可来自共轭化合物如酮或芳香化合物如三甲苯等，称 π 给予体。接受体一般为缺电子化合物，如三硝基苯等，分子内若存有部分正电荷，就易于接受电子。

鉴于量子化学分子轨道计算可获得分子轨道能级，而在电荷转移复合物形成时，给予的电子一般自最高占有轨道(HOMO)跃迁，而接受的电子一般安放在最低空轨道(LUMO)。这样依据量子化学计算所获的最高占有轨道能量(E_{HOMO})和最低空轨道能量(E_{LUMO})，即可估计电荷转移是否易于产生。如给予体中，E_{HOMO} 愈高，愈易于给出电子。接受体中，若 E_{LUMO} 愈低，接受的电子也愈稳定。

早在 50 年代，就有不少人用简单的分子轨道方法(HMO 法) 计算了大量药物分子。在对致幻剂的研究中，发现色胺类化合物的致幻作用与经 HMO 计算获得的 E_{HOMO} 值平行，这是量子药理学早期的研究成果之一。此外还通过 HMO 法计算，推测了治疗精神分裂症的药物氯丙嗪是个极好的电子给予体。还有的曾用 E_{HOMO} 和 E_{LUMO} 能级讨论了抗疟药物的活性。研究表明，一些生物碱如烟碱、阿托品、毒扁豆碱、吗啡、士的宁等都易于给出电子。

随着量子化学计算方法的广泛应用，半经验方法尤其是从头计算法用于计算药物分子以来，探索分子轨道能级，从 E_{HOMO} 和 E_{LUMO} 的大小中寻求电荷转移复合的形成，往往与药物构象、分子静电势等紧密配合、相辅相成，从而推动了更深入地探讨药物与受体作用的研究。

值得注意的是，在探讨电荷转移复合物形成过程中，除了注视 E_{HOMO} 和 E_{LUMO} 之外，还要同时分析其他临近 LUMO 的空轨道能量。

在探索抗疟药物青蒿素及类似物的受体作用研究中[64]，各衍生物的原子电荷分布几乎保持不变，最引人注目的 E_{HOMO}、E_{LUMO} 能量（由 EHMO 计算获得）几乎相同，然而第四空轨道能级却表现了明显的差别。若将生物活性数据 log 1/C（C 表示抑制疟原虫 90% 所需药物克分子浓度)与第四空轨道能级 E_{FUMO} 作图（图 8-83)，呈良好的线性关系。经回归计算，所得相关方程，其相关系数 r 达 0.99(见式 8.8)

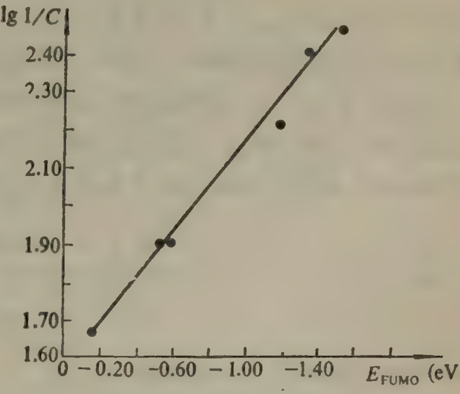


图 8-83　log 1/C～E_{FUMO}相关图

$$\log 1/C = 1.573 - 0.592\,E_{FUMO} \tag{8.8}$$

$$n=6;\ r=0.994;\ s=0.0402$$

这个现象已经引起不少学者的重视[83,84]，其中 Kikuchi 阐明当生物活性分子与受体 DNA 作用时，可近似地看作是与 DNA 中鸟嘌呤上氨基相作用。并指出，鸟嘌呤的氨基是个外层有 8 个电子的原子团，它与生物活性分子形成电荷转移复合物时，分别依次将 8 个电子填入对方分子轨道中的空轨道。经分析，鸟嘌呤上的氨基属于苯胺型N—C键，氨基对于苯环采取若干扭曲的结构，因而 3 个2 p 轨道分别与碳原子及二个氢原子形成三个 σ 键，而氮原子上的 2 个 s 电子则

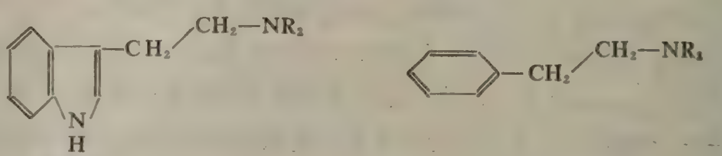
处于孤对电子状态[85]。由此可见，填入第四层空轨道的是孤对电子。因而亦可将第四层空轨道能级高低作为与受体作用的理论描述。

从青蒿素衍生物的能级来看，虽然 E_{LUMO} 很接近，但 E_{FUMO} 有较大差别，E_{FUMO} 高，则鸟嘌呤氮上孤对电子不易填入，结合就差，亦即抗疟活性低。E_{FUMO} 低，则孤对电子易填入，形成电荷转移复合物，亦即表现较高抗疟活性。

这就告诉人们，在探讨电荷转移时，尤其是当受体为核酸时，除了考察 E_{HOMO} 和 E_{LUMO} 外，还应注视其余临近的轨道，特别是第四层空轨道能量。

8.5.2 致幻剂的电荷转移研究

致幻剂能对人类主观精神作用产生显著的影响。色胺和苯乙胺是具有致幻作用的两种主要的分子类型。Snyder 和 Merril 利用 HMO 法对一系列上述分子类型的化合物进行了计算[44]，他们比较了 E_{HOMO}、E_{LUMO}、π 电荷、前沿电子密度、超离域度等分子轨道指数与致幻活性之间的关系，发现化合物 E_{HOMO} 值与致幻活性有良好的对应关系。



色胺类型　　　　苯乙胺类型

表 8-12 列举了三个 α-甲基苯乙胺的计算值，显示了 E_{HOMO} 与致幻活性之间的良好平行关系，表中的致幻活性是相对于三甲基苯乙胺的活性（以下各表同理）。一系列苯乙胺类似物的 HMO 计算列于表 8-13，表明在苯乙胺类型中，墨斯卡灵（三甲氧基苯乙胺）的 E_{HOMO} 最高，应具有最强的致幻活性，而表中其他分子没有致幻活性，从表 8-12 和表 8-13 数据可知，最高占有轨道能量必须具备相应的高能量才具有致幻活性。

表 8-12　苯乙胺类致幻剂活性与 E_{HOMO}

化　合　物	致 幻 活 性	$E_{HOMO}(\beta)$
3,4,5-三甲氧基-α-甲基苯乙胺	17	0.4310
2,4,5-三甲氧基-α-甲基苯乙胺	2	0.5357
2,3,4-三甲氧基-α-甲基苯乙胺	<2	0.5696

表 8-14 列出了另一类型致幻剂的计算结果，可见色胺衍生物与苯乙胺类情况一样，E_{HOMO} 越高，致幻活性愈强。

经研究，迄今最强的致幻剂麦角酸二乙酰胺(LSD)，其 $E_{HOMO}=0.218\beta$，这远远大于表8-15所列化合物的 E_{HOMO}。

作者为进一步讨论 E_{HOMO} 与致幻活性关系，比较了 5 个具有致幻作用而结构不同的化合物，发现 E_{HOMO} 与致幻力亦成正比关系。（表 8-15）

表 8-13 苯乙胺类(无致幻活性)的 E_{HOMO}

化合物	名称	$E_{HOMO}(\beta)$
3,4,5-三甲氧基	三甲氧基苯乙胺	0.5357
2,3,4-三甲氧基		0.5696
3,4-二甲氧基		0.5702
4-甲氧基	甲氧基酪氨	0.6583
3-甲氧基	甲氧基间酪氨	0.7240
3,4-二羟基	多巴胺	0.6586
4-羟基	酪氨	0.7209
—	苯乙胺	0.8619

表 8-14 一系列色胺衍生物 E_{HOMO} 与致幻活性关系

化合物	名称	致幻力*	$E_{HOMO}(\beta)$
4-羟基-N,N-二甲基	Psilocin	31	0.4603
6-羟基-N,N-二甲基		25	0.4700
5-羟基-N,N-二甲基	Bufotenin	弱	0.5147
N,N-二乙基		弱	0.5164

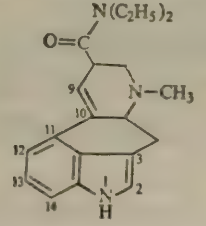
* 与三甲氧基苯乙胺在有效剂量时的比值。

表 8-15 不同种类药物中与 E_{HOMO} 相对应的致幻活性

化合物	最小有效剂量(毫克/公斤)	活性*	$E_{HOMO}(\beta)$
麦角酸二乙酰胺(LSD)	0.001	3700	0.2180
Psilocin	0.12	31	0.4603
6-羟基-二乙基色胺	0.15	25	0.4700
TMA-2	0.22	17	0.4810
TMA	1.70	2.2	0.5357
三甲氧基苯乙胺	3.75	1	0.5357

* 比三甲氧基苯乙胺有效剂量之比值。

对 LSD 分子的深入分析,认为 LSD 分子中 C 9、C 10 双键与吲哚共轭,于是提高了 HOMO 能级,成为电子给予体,实验证明,将此双键饱和,氢化、溴化或加 H_2O, 则致幻作用剧降。 LSD 的C 2 位有较高电子密度,可能在 2 位形成电荷转移复合物。



麦角酸二乙酰胺(LSD)

上述研究表明，致幻活性的产生是由于致幻剂乃是电子给予体，其与受体形成电荷转移复合物。

8.5.3 抗心律失常新药常咯啉及其类似物的量子药理学研究

常咯啉是我国自行创制的抗心律失常新药。动物试验表明它对乌头碱引起的动物心律失常具有明显的防治作用。临床试验报告总有效率达 80.8%，对阵发性室性心动过速和频发性室性早搏疗效尤为显著。在对常咯啉药理作用继续深入研究的同时，合成了若干改进的化学结构类似物，在探讨了它们的构效关系之后，吴吉安等[18]用量子化学 EHMO 方法对常咯啉分子及一系列类似物进行了计算。

首先对常咯啉分子进行能量优选，以获得稳定构象（详见 8.2）。最稳定构象视作常咯啉的"活性构象"。图 8-84 显示了常咯啉分子用"活性构象"计算所获的分子电子图，括弧内为键级。

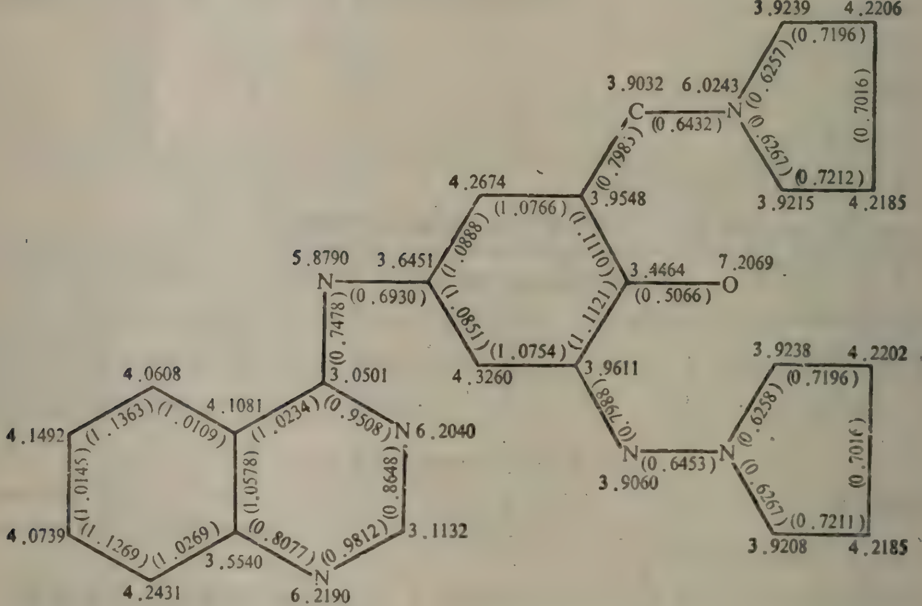


图 8-84 常咯啉分子电子分布图

选用已经合成并有定量药理数据的常咯啉类似物 5 个，这些化合物的共同特点是以苯甲酰基代替常咯啉的喹唑啉环，结构如图 8-85 所示。这些类似物结构与常咯啉相似，苯甲酰基中苯与羰基因 $\pi-\pi$ 共轭而处于同一平面，而C＝O与常咯啉的喹唑啉中C＝N相近，故而类似物的计算仍按常咯啉的优选后构象处理。

5 个类似物的 EHMO 计算结果列于表 8-16，将表中所列的分子轨道指数以及未列出的如电荷密度、键级等指数与药理试验生物活性进行广泛的相关分析计算，发现除了 E_{LUMO}、E_{FUMO}、

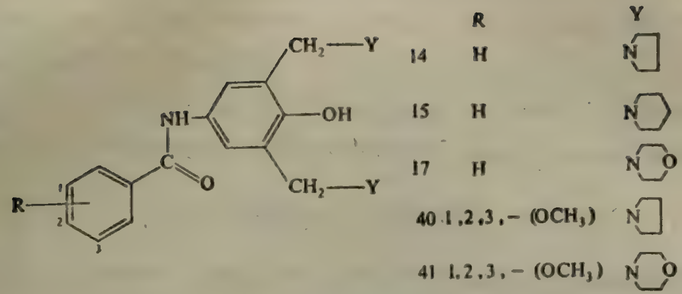


图 8-85　常咯啉类似物分子结构

表 8-16　常咯啉类似物的EHMO计算

分子编号	14	15	17	40	41
分子式	$C_{23}H_{29}N_3O_2$	$C_{25}H_{33}N_3O_2$	$C_{23}H_{29}N_3O_4$	$C_{28}H_{35}N_3O_5$	$C_{26}H_{35}N_3O_7$
原子数	57	63	59	69	71
价电子数	148	160	160	184	196
轨道数	141	153	149	171	179
占有轨道数	74	80	80	92	98
E_{HOMO}	−11.53149	−11.41260	−11.47271	−11.57162	−11.49588
E_{LUMO}	−9.76178	−9.76241	−9.76103	−9.72518	−9.72437
E_{LOMO}	−37.38043	−37.32629	−37.37929	−37.71910	−37.71854
$E_{HOMO}-E_{LOMO}$	25.84894	25.91379	25.90658	26.14748	26.22266

注：所有能量均为电子伏特。E_{HOMO}为最高占有轨道能量，E_{LOMO}为最低占有轨道能量。

表 8-17　生物活性数据与 E_{HOMO}-E_{LOMO} 指数间的相关方程

生物活性数据		a	b	r	s	t	eq.
大鼠BA有效剂量		12.92	−5.68	0.969	0.117	−6.79	(1)
大鼠BA致死剂量		7.43	−3.60	0.979	0.061	−8.28	(2)
小鼠iv的 LD_{50}		9.96	−4.64	0.916	0.164	−3.95	(3)
药物对抗BA引起的室性早搏的作用	小鼠ED_{50}	13.88	−6.32	0.966	0.136	−6.51	(4)
	大鼠ED_{50}	15.70	−7.05	0.9[illegible]1	0.205	−4.80	(5)

相关方程为 $\log 1/C=a+b.(E_{HOMO}-E_{LOMO})$，$r$为相关系数，$s$为标准偏差，$t$为系数 t 检验值，BA 为北草乌头碱。

E_{LOMO}、$E_{HOMO}-E_{LOMO}$ 等能量值外，其余指数与生物活性均无明显线性关系。其中尤以 $E_{HOMO}-E_{LOMO}$ 指数相关性最好。见表 8-17 中方程(1)—(5)。

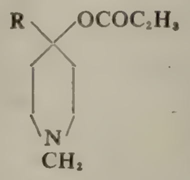
从上述计算分析，生物活性与轨道能级有显著相关，其中又以 $E_{HOMO}-E_{LUMO}$ 线性最佳。这项指数与目前所获的药理数据都导得较好的相关方程($r>0.91$)。这就提示我们，常咯啉在体内可能的作用形式是电荷转移。实验研究表明，常咯啉抑制慢反应 AP，并不使 ERP 缩短．相反使 APD 和 ERP 明显延长，由于 f 阀门闭合延缓，产生强大的 Ca^{2+}内流，认为常咯啉的抗心律失常

可能是一种新的类型，主要是阻断兴奋折返，细胞中聚集较多 Ca^{2+}，引起抑制心律失常作用。从表 8-16 方程(1)—(5)可知 $E_{HOMO}-E_{LOMO}$ 值愈小，生物活性作用愈大，这一指标一定程度上标明失去电子能力的大小，$E_{HOMO}-E_{LOMO}$ 小，易给出电子，反之亦然。这样易给出电子，就易与金属离子如 Ca^{2+}结合，也就利于较多的 Ca^{2+}内流，这似乎与实验事实相吻合。这些结果有助于设计活性更高的化合物(如 $E_{HOMO}-E_{LOMO}$ 更小)，并对推测作用机制提供了有益的线索。

8.5.4 镇痛药中的电荷转移研究

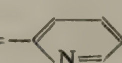
度冷丁自成为新型结构的鸦片受体镇痛剂以来，对其进行了结构改造，合成了一系列类似物，其中普鲁丁是较为成功的度冷丁型结构的镇痛药物。

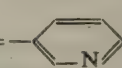
对普鲁丁类似物的构效关系研究表明，电荷转移复合物的形成对其与受体的相互作用起了至关重要的影响[86]。

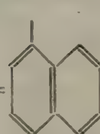


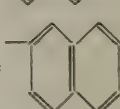
R =  普鲁丁

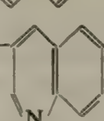
I：R =  3-thieny

II：R =  2-pyridyl

X：R =  3-pyridyl

XI：R =  1-naphthyl

XII：R =  2-naphthyl

XIII：R =  3-quinolyl

经考察上述 6 个普鲁丁类似化合物的生物活性与 π 轨道 E_{LUMO} 和 E_{HOMO} 值，发现存在良好的线性关系。

表 8-17 列出 6 个化合物的生物活性与 E_{HOMO} 和 E_{LUMO} 值，ED_{50} 为热板法 (hot plate method)测得。π 轨道能量采用 HMO 法计算各自的共轭的 R 取代基获得(事实上作者用有关手册的计算结果)。

表 8-18 普鲁丁类似物的生物活性与 E_{HOMO} E_{LUMO} 值

化合物	ED_{50} mg/kg SC	E_{HOMO} (β)	E_{LUMO} (β)
I	3.90	0.2928932	−1.00
II	16.00	0.5602315	−0.5602315
X	16.60	0.5602315	−0.5602315
XI	22.00	0.6180340	−0.6180340
XII	22.50	0.6180340	−0.6180340
XIII	20.00	0.7032833	−0.5270970

从表中可知，随着 E_{HOMO} 值的减少，其镇痛活性随之下降，而 E_{LUMO} 也随着降低。（注意在 HMO 计算中，能级以 β 为单位，其间与通常能量相差一个负号）。图 8-91 清晰地表明了 ED_{50} 与 E_{HOMO} 之间的线性关系。

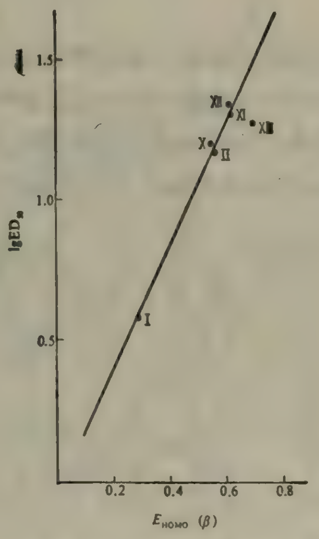
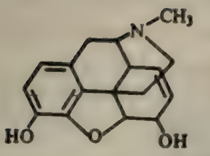


图 8-86 E_{HOMO} 值与 ED_{50} 相关图

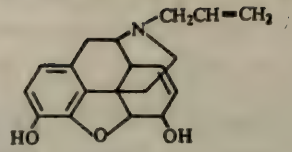
上述研究表明，在普鲁丁类似物的受体作用中，电荷转移复合的产生起着重要作用。普鲁丁中的苯基起电子给予体的作用，若换以其他基团，如相应的噻吩基，也是电子给予体，仍有相近的镇痛作用。所计算的 6 个化合物中，噻吩 E_{HOMO} 最高，很好地说明了这一点。它仍具有较强镇痛作用。吡啶基化合物（II、X）E_{HOMO} 较低，镇痛作用就较弱。如换成萘基或喹啉基，它们的 E_{HOMO} 更低，这些化合物的镇痛活性也就更小。

8.5.5 吗啡类药物的电荷转移复合研究

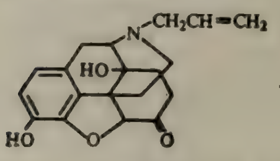
众所周知，吗啡生物碱的镇痛作用非常显著，同时它还有抑制呼吸、眩晕、恶心、呕吐、便秘等副作用。在寻找高效低毒的镇痛药中，药物学家作了广泛而深入的探索。若将吗啡氮上的取代甲基换成烯丙基，成为丙烯吗啡(nalorphine)，却产生了意外的药效，它不仅会产生对抗吗啡的镇痛作用，而且也抑制了后者的呼吸抑制、便秘、呕吐、缩瞳等作用，所以丙烯吗啡成为治疗鸦片中毒及由于服用吗啡类药物的呼吸抑制作用的药物，若进一步修饰结构为纳洛酮(naloxone)，则是更强的吗啡对抗药物。



吗啡



丙烯吗啡



纳洛酮

为什么结构上的微小变异便导致对抗作用呢？这自然引起量子药理学家的很大兴趣。

Cheney 等[87]用量子化学方法对吗啡和纳洛酮作了较深入的研究，他们首先使用分子力学方法分别计算了纳洛酮分子自由基与 N 质子化的构象，计算表明吗啡和纳洛酮两者构象相似。图 8-87 表示了纳洛酮分子的构象能图。

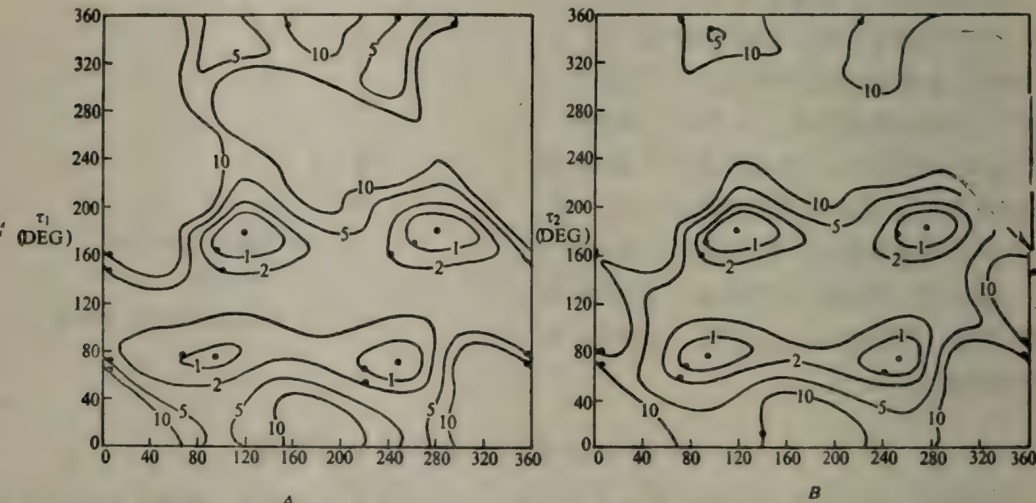


图 8-87 纳洛酮分子 *A*.自由基，*B*.N 质子化构象能图(kcal/mol)

进一步使用量子化学从头计算中的浮动球 Gaussian 轨道法(floating spherical Gaussian orbital method 简称 FSGO 法)对吗啡和纳洛酮分子各自的自由基和 N 质子化进行了分子轨道计算，除获得能量指数外，还进行了分子静电势计算，图 8-88 至图 8-89 分别是吗啡和纳洛酮分子静电势图。

从静电势图分析，吗啡与纳洛酮分子静电势正、负区域相似，可见从构象分析及静电势比较均不能找到这两个分子药理作用相对抗的原因。

对计算结果的分析表明，主要差别在于纳洛酮阳离子的最低空轨道能量 E_{LUMO} 低于吗啡

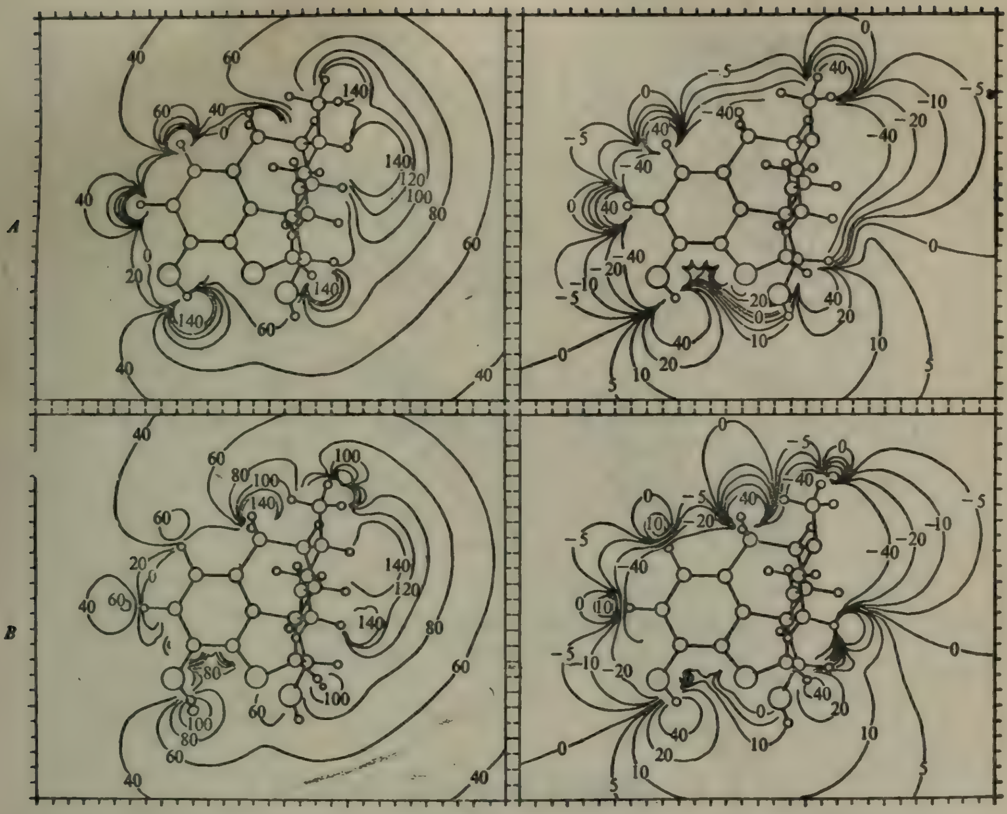


图 8-88 吗啡分子静电势图。左为自由基，右为质子化(下同)
A. $z=0$ 的 xy 平面，*B*. $z=1.36$a. u.的 xy 平面

阳离子的 E_{LUMO}，图 8-90 显示了两个分子的能级图。

从图 8-90 中可知，吗啡阳离子 $E_{\text{LUMO}}=0.218$a. u. 而纳洛酮阳离子 $E_{\text{LUMO}}=0.1698$a. u.，后者明显低于前者。经分析 LUMO 轨道波函数可知，两个分子的 LUMO 均属于 π 型。这一事实证明丙烯吗啡或纳洛酮中的烯丙基可作为电子接受体，而吗啡氮上的甲基则不易接受电子，亦即前者可与受体部位形成电荷转移复合物，而后者却不易形成。

于是，根据推测，烯丙基对抗物的受体作用图应如图 8-91 所示。X、Y、Z 是受体上三个电负性部位，其中 X 与苯环大 π 键产生电荷转移复合物，Y 与羟基的氢原子产生氢键作用，Z 可与吗啡阳离子产生静电引力。吗啡的受体作用主要由这三个部位产生。在烯丙基引入的情况下，受体转变部分构象，Z 移至图中虚线所示部位，与烯丙基产生电荷转移作用，亦即药物由于与受体的电荷转移作用改变了受体构象，乃把激动作用改变为对抗作用。

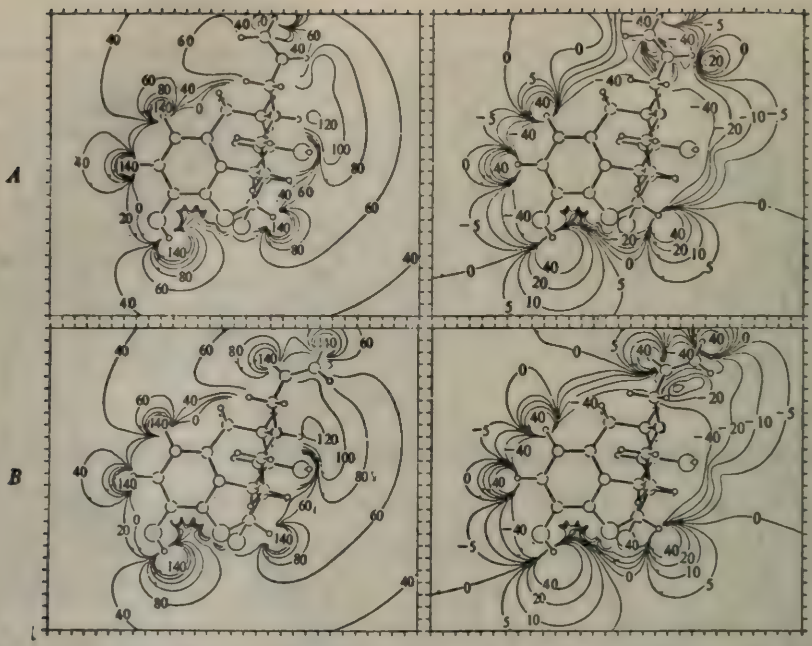


图 8-89 纳洛酮分子中 $z=0$ 的 xy 平面静电势图。
A 为$\tau_1\tau_2$(180.5, 121.4)B 为 $\tau_1\tau_2$(180.5, 275.0)

8.6 药物分子设计与 QSAR

8.6.1 引　　言

如何开发高效低毒的新药，是每个药物化学和药理学研究工作者的共同愿望。自从上一世纪起至近代，这个愿望大致经历了三个发展阶段。第一阶段由远古至上世纪末是发现药物时期(Discover)，如植物药、生物碱等。第二阶段由本世纪初至五十年代是药物开发时期(Develop)，如合成药物的出现，典型的例子是磺胺系列药物。第三阶段由六十年代初起进入药物设计时期(Design)，定量药物设计的产生，多种学科的相互渗透配合。

从当今寻找新药的途径看，不外乎以下五个方面：

(1) 由天然产物，如动物、植物、矿物、海洋生物、微生物中分离提取有效成分，并以此作为先导化合物(lead compound)，进行结构改造，合成并筛选一系列衍生物或类似物，从中寻找出更有效的毒性更低的新药。如我国从中草药蒿草中分离得青蒿素，是一种新型结构的抗疟药物，经合成筛选，找到了抗疟新药——蒿甲醚。

(2) 随机筛选。“泛泛合成，随机筛选”，方法较古老，带有较大偶然性。

(3) 现有药物的结构改造。如抗癌药物氮氛类化合物的结构修饰。

(4) 研究已知药物的体内代谢过程，发现其活性代谢产物，从而引导设计新药。

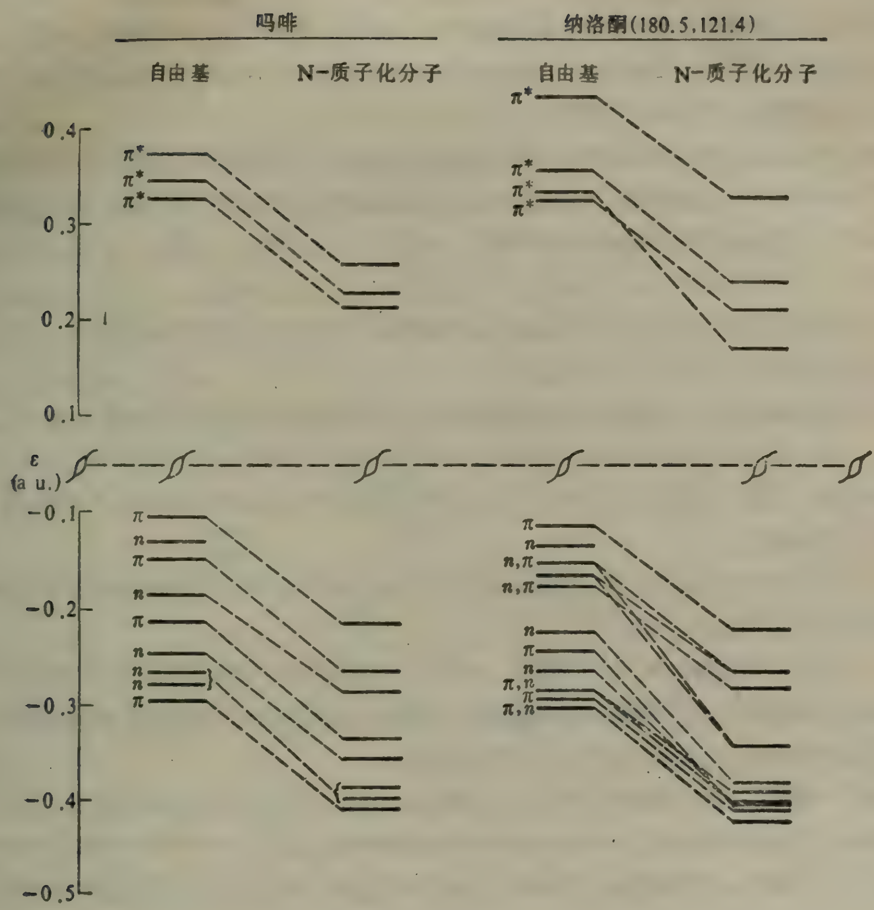


图 8-90 吗啡与纳洛酮的自由基与 N 质子化分子能级图

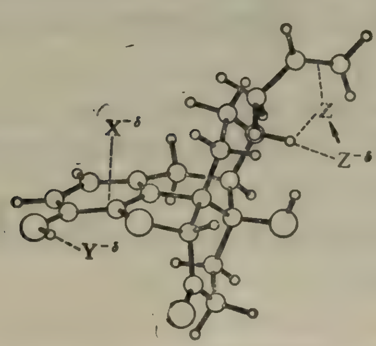


图 8-91 纳洛酮 N 质子化分子受体作用图

(5) 在药理作用指导下设计新药。

在本世纪六十年代，由于电子计算机的广泛使用，随着数学、物理、药理、生物化学和药物化学等诸领域的互相渗透，作为药物分子设计这一药物研究的中心内容逐渐由定性转向定量。于是药物的定量构-效关系(quantitative structure-activity relationships 简称 QSAR) 的研究应运而生，并逐渐获得广泛应用[88-91]。

定量构效关系是用计算机对现有活性化合物(如一系列相同类型药理作用的同源化合物)的信息进行数学解析，用数学的模式描述有机化合物分子结构的理化参数或结构参数与生物大分子化合物(如酶、辅酶)或生物大分子组织(如 DNA，血清蛋白，细胞，动物) 之间相互作用关系。

QSAR 研究使得药物从开发至临床应用的周期大为缩短，效率显著提高。在 QSAR 研究中，药物分子的设计通常可分四步进行。

(1) 用前述的第一个寻找药物途径的方法，选择有发展前景的先导化合物。

(2) 合成并筛选先导化合物的衍生物。QSAR 研究力图使这个过程从无止境地合成缩短到较合理的程度。实践表明，往往用 8—10 个衍生物就可获得化合物的理化性质和药理生物效应之间相互关系的概念。

(3) 建立一个初步的定量构效关系方程式，依据方程式提供的药物设计指示信息，进一步设计并试验衍生物，从而开拓 QSAR。

(4) 建立一个理想的 QSAR 方程式，且所有与生物活性相关的理化性质和结构参数的最佳值也已确定。这时即可向"最佳"衍生物的合成迈进。当然，这一步多少带有假设性质，而预测化合物的活性最终是否令人满意还需由药理试验决定。

在 QSAR 研究中，应用最早，使用最广的是由 Hansch 等 1964 年建立和发展起来的超热力学法(或称线性自由能加和法，亦称 QSAR 中的 Hansch 途径) (extrathermodynamic 或 linear free energy)。本节主要介绍该法的基础、参数和应用，最后简略地介绍 QSAR 中的其他方法。

8.6.2 QSAR 中 Hansch 途径的基础

Hansch 途径的数学模型是建立在化合物生物活性与其理化性质直接相关的基础上的。依据 Hammett 方程式，采取与自由能有关的取代基物理化学参数的线性组合来描绘化合物的结构与生物活性之间的定量关系式

$$\log 1/C = a\pi + b\sigma + cE_s + d \tag{8.9}$$

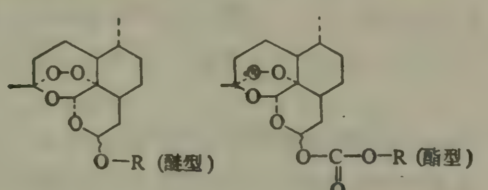
其中，π 为化合物取代基的疏水常数，σ 为 Hammett 取代基电子效应常数，E_s 为 Taft 取代基立体效应常数。

研究表明，决定药物生物活性强弱的主要因素取决于药物在其作用部位(即受体) 表面的浓度和药物与受体的亲和力与内在活性。

Hansch 认为，药物到达受体表面浓度的高低与药物在体内的转运(transport)相关。而药物的吸收、分布与代谢又与其本身的理化性质有密切关系。他进一步指出，药物在体内的随机转

运(如透过细胞膜或血脑屏障)与化合物油水分配系数 $\log P$ 或代表其疏水亲脂作用的取代基疏水常数 π 密切相关。同时,药物的内在活性和受体间的亲和力也与药物本身的理化性质有密切的关系。一般说来,化合物-受体相互作用可涉及电子效应(如离子键、离子-偶极、偶极-偶极相互作用),疏水相互作用(如形成疏水性或亲水性囊),次级力相互作用(诱导偶极-偶极相互作用、氢键、Van der Waals 引力等),以及电荷转移相互作用(π-π 相互作用)。因此,结构相似的同源化合物系列中,其生物活性可用化合物的物理化学参数来描述。如方程(8.9)所列,它应包括三方面的参数:表示药物在体内随机转运作用的油水分配系数 $\log P$ 或疏水常数 π;用 σ 和与 σ 有关的其他电性参数表示药物与受体的作用;用 E_s、MR(摩尔折射率)等描述空间作用的立体参数表示药物作用和基团大小对靶分子的影响。这种线性自由能加和法既有一定的物理意义,又紧密地构筑起药理作用与化合物理化性质的联系,这就是 Hansch 途径的特点,也是它受到广泛重视和普遍应用的根本原因。

以下我们以抗疟新药青蒿素衍生物的 QSAR 研究为例,说明 Hansch 途径建立的过程[92]。



通过 Hansch 途径的研究,在化合物生物活性与其物理化学参数之间已经建立了如下方程(称相关方程)

$$\log 1/C = 3.112 - 0.142(\pm 0.036)\log P - 0.559 I_{\alpha,\beta} \tag{8.10}$$

$$n=49;\ r=0.829;\ s=0.205$$

$$\log 1/C = 1.120 - 0.198(\pm 0.033)(\log P)^2 + 1.155(\pm 0.215)\log P - 0.612(\pm 0.048) I_{\alpha,\beta} \tag{8.11}$$

$$n=49;\ r=0.911;\ s=0.153;\ F=72.8;\ \log P_0=2.91$$

$$\log 1/C = 1.945 + 0.347(\pm 0.08)\log P - 0.991(\pm 0.155)\log(\beta P+1) - 0.615(\pm 0.05) I_{\alpha,\beta} \tag{8.12}$$

$$n=49;\ r=0.915;\ s=0.150;\ F=76.8;\ \log P_0=3.03;\ -\log\beta=3.29$$

其中 $\log 1/C$ 代表化合物的定量生物活性,C 是产生生物活性的药物摩尔浓度。生物活性通常以 ED_{50}, LD_{50}, MIC 等为指标。在同一系列化合物中,$\log 1/C$ 必须在相差 1 个数量单位以上的范围方能导得具有指导意义的相关方程。在方程(8.10)至(8.12)中的生物活性 C 为 SD_{90},系杀灭疟原虫 90% 的药物量。$\log 1/C$ 经摩尔校正。$\log P$ 是化合物油水分配系数,由实验测定。$I_{\alpha,\beta}$ 为指示变量(indicator variable),表明当青蒿素衍生物为 α 构型时 $I_{\alpha,\beta}=0$,β 构型时 $I_{\alpha,\beta}=1$。

Hansch 途径中相关方程的系数及统计学指数往往要用基于最小二乘法的多元线性回归方

法求得。目前已有不少专用于 QSAR 研究的计算机程序发表。大多数研究者一般通过自编程序在小型或微型计算机上完成。

在方程(8.10)至(8.12)中，n 为化合物数。一般情况下，为保证统计显著性和相关方程的可靠性，若参数个数为 m，则要求 $n \geq 5m$。r 为相关系数(correlat on coefficient)，表明相关程度，以 r^2 为标准。方程(8.10)中 $r^2=0.83$，说明该方程拟合了 83%的生物活性数据。s 为标准偏差(standard deviation)，F 为 F 检验值(F-test)。

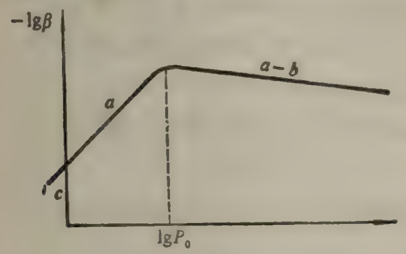


图 8-92 双线模型曲线图

在 Hansch途径的研究中，尤其关注药物在体内的转运过程；往往油水分配系数是首选和最重要的参数。通常在活性与油水分配系数之间并没有简单的直线关系。Hansch 通过大量的深入研究，发现数百类化合物的 $\log 1/C$ 与 $\log P$ 呈抛物线关系，即引入$(\log P)^2$ 项后，方程相关性显著提高。(见方程(8.11))。钟罩形抛物线顶点对应的 $\log P_0$ 系活性最佳的 $\log P$ 值。

随着 QSAR 研究的逐步深入，发现很多药物的生物活性与 $\log P$ 之间的关系，并不象抛物线这样对称，而是由两条歪斜的上升和下降直线通过一小段抛物线连接而成的双线性图形（图 8-92)。

Kubinyi[93] 依据 McFanland 理论提出了 Hansch 途径中的双线模型 (bilinear model)，方程(8.12)即为典型的双线性方程，如图 8-92 所示。曲线的上升部分和下降部分的两条直线，其斜率分别为 a 与 $a-b$。它们的连接处呈抛物线，活性最大值所对应的 $\log P$ 为最适$(\log P)_0$，C 为截距。

通过对近百类化合物的构-效关系的回顾性考察看到，在大多数情况下，双线模型的各种统计学指标均比抛物线模型好，用它拟合的方程更符合药物构-效关系的实际情况。方程 (8.12) 显然比(8.11)要好。吴吉安等曾对国内发表的抛物线方程用双线方程拟合，获得了同样的结论[94]。

8.6.3 QSAR 研究中的参数

化合物的油水分配系数 $\log P$ 被认为是 Hansch 途径中最重要的参数。

生化系统的分配系数 $P_{生物相}$ 指化合物在水相和产生标准生物效应的临界生物相之间的分配系数，其定义为

$$P_{生物相}=\frac{C_{生物相}}{C_{水相}} \tag{8.13}$$

$C_{生物相}$ 和 $C_{水相}$ 分别为药物在生物相和水相中的摩尔浓度。因 $P_{生物相}$ 不易测定，不得不用正辛醇-水系统来模拟。据实验证明，该系统为良好的参考系统。$\log P$ 和疏水常数 π 都是可加和性的自由能项

$$\log P = \sum_{i}^{n} \pi_i \tag{8.14}$$

可将分子拆成 n 组分，$\log P$ 可通过加和分子各组分的 π 值获得，同时，同源物的 $\log P_{RX}$ 值也可通过下式求得

$$\log P_{RX} = \log P_{RH} + \pi_X \tag{8.15}$$

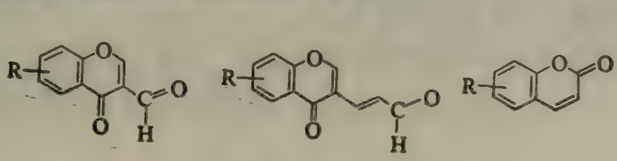
当母体的 $\log P_{RH}$ 值为已知或通过实验测定的情况下，且有取代基常数 π_X 可利用，如果在取代基与母体已有的基团无相互作用或这种作用预先可测定的情况下，计算的 $\log P_{RX}$ 值与实验值将十分接近。后来 Hansch[95]与 Rekker[96]等分别提出了 $\log P$ 的分子碎片计算法 (fragment method)是目前常用的计算方法，详细介绍可参阅有关文献。

当今各国学者仍注重于实验测定 $\log P$ 的研究，沿用传统的摇瓶法，用 UV 分光光度计测定经分配和离心后在水相或正辛醇中的药物浓度。Hansch 所在的 Pomona 学院药物化学系正在源源不断地把实验测得的新的 $\log P$ 值输入其数据库[97]。他们又以专著的形式列出了 3 万多个化合物 $\log P$ 实验值及众多的各类取代基参数[95]，这已成为QSAR 研究的必备手册。由于以溶剂分配方法测定 $\log P$ 相当繁杂，新近的研究致力于色谱法测定 $\log P$ 值，其中尤以高效液相色谱(HPLC)以其快速、易测、重现性好等优点，逐步得到推广[98]。

代表电子效应的参数是大量的，但基本上都可以从文献查到。最广泛使用的参数包括：Hammett 常数 σ 的各种变型，描述取代于间位和对位侧链上的取代基电子效应常数 σ_m，σ_p，以及用于脂肪族系统的 Taft 极性取代基常数 σ^*。如苯环上取代基的对位原有强烈的推电子或吸电子基团存在，则会产生超共轭(hyper conjugation)或共轭通路(through resonace) 效应，则须用 σ 或 σ^+ 来代替相应的 σ。在酸性化合物系列中，电离常数 pK_a 往往是重要参数。此外有的还用诱导效应常数 F 和共振效应常数 R。

除了上述电性参数外，量子化学计算得到的各种分子轨道指数，也常常用于 Hansch 途径的研究中，事实上，读者已从本章的不少例子中看到量子化学计算在QSAR中的广泛应用。

值得指出的是，量子化学计算中，对于不同母核结构的分子中的共同特征的部分，可将这部分(如原子或基团及静电势区域等)的量子化学指数作为参数与同一生物活性进行相关分析，这就不仅限于同源化合物的QSAR，从而扩大了 Hansch 途径的适用范围。如在一系列含 α、β 不饱和酮基结构的色酮、萘醌类、香豆素的平喘活性研究中导出了平喘活性与羰基及邻近的 α、β 不饱和双键碳原子的分子轨道指数之间的相关方程[99]。



$$\log \Delta t = -10.0 + 8.69\, q_{C=O} \tag{8.16}$$

$$n = 14;\ r = 0.92;\ s = 0.166$$

$$\log \Delta t = 0.825 + 7.72\, f_{C=O} \tag{8.17}$$

$$n=14;r=0.892;s=0.19$$

式中 $q_{C=O}$ 是由 HMO 法计算所获，表示羰基碳氧原子电荷密度之和。$f_{C=O}$ 是羰基两个原子前沿电荷密度之和，$\log\Delta t$ 系平喘活性。

空间效应参数一般都可以从文献中查得[95]。早期使用较广的空间参数是Taft立体参数 E_s，后来经 Hansch 等推荐较普遍采用的描述取代基体积大小的参数摩尔折射率 MR。近年来，由 Verloop 提出的描绘取代基长度及不同方向宽度的 STERIMOL 参数 L, B_1、B_2、B_3、B_4、B_5 得到了较普遍的应用，Fujita 等推崇的 Van der Waals 体积差 ΔV_w 参数也获得了较好的应用。

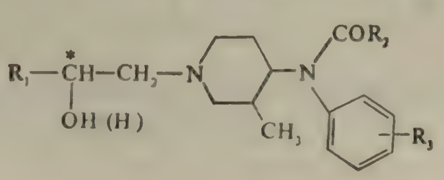
除了上述描绘取代基疏水性、电子性、立体性的物理化学参数外，还选取了作为不连续变量的结构因素参数。这些参数的选用，时常取决于研究者对QSAR 对象的分析。它与特定的原子、键、环、官能团、成对的官能团、其他特定的结构部分及立体构型的有无，或者还有某些不明确的因素密切相连，这些参数以 0—1 型变量表示，称为指示变量，它时常在 Hansch 途径的研究中起到举足轻重的作用。

在选择各种参数进行相关方程的计算时，所选取的不同种类的参数之间应相互独立，以避免依赖关系(这称为参数间的共线性(corelinear))存在。如许多取代基的 π 与 MR 常常有共线性，若选用这些取代基来改变化合物的结构，就无法分辨生物活性的改变究竟由疏水性不同还是由立体性差异(MR)所致，从而还须合成更多的新化合物。因而，从 QSAR 角度出发，设计化合物进行合成时，各理化参数变化幅度尽可能大，且以取代基的各种参数间没有明显相关的取代基为首选合成，这样方能获得有意义的结果。在这方面，Hansch 等曾建议采用类聚群分析 (cluster alalysis)方法，有明显的指导作用。

8.6.4 QSAR 研究的应用

(1) 预测同源物的生物活性，指导药物的设计与合成。当我们对已合成并获得了定量活性值的一系列同源化合物通过 Hansch 途径获得较好的相关方程以后，即可根据方程中参数情况、系数的正负、抛物线或双线性的最适值、指示变量等进一步选择新取代基，并在合成前事先用方程计算出可能的活性，这样既可找到活性更高的化合物，又避免了盲目合成。如方程(8.9)至(8.10)中预示的青蒿素醚类酯类衍生物的相关方程中，指出($\log P_0$) = 3.0 左右，经考证已合成的化合物，已有几个化合物的 $\log P$ 值在 3.0 附近。这提示，在这种类型的衍生物中，由于抗疟活性仅与油水分配系数相关，可能已经找到活性最适的化合物，下一步主要应开发新的衍生物类型的合成，这一信息，给合成工作提供了有意义的参考。

强效镇痛剂 3-甲基芬太尼衍生物的构效关系研究中，导得不同位置取代基理化参数与镇痛活性 ED_{50}之间的相关方程[100]，当 R_1、R_2、R_3 取代基变化时，经采取逐步回归分析法导得方程(8.18)



$$\log 1/C = 6.768 + 0.136\,I_0 + 0.340\,MR_{R_2} - 0.313\,MR_{R_2}^2 + 1.451\,\sigma - 17.091\,\sigma^2 \qquad (8.18)$$

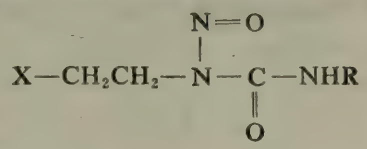
$$n = 39;\ r = 0.874;\ s = 0.498;\ (MR_{R_2})_0 = 0.543$$

$$\sigma_0 = 0.042$$

式中 MR_{R_2} 代表 R_2 取代基摩尔折射率的 1/10，σ 为 R_3 取代基电子参数，I_0 系指示变量，当 C* 位置有 OH 取代时取为 1。进一步合成选取 R_2 和 R_3 应在方程预示的最佳值附近，于是 R_2＝—OCH_3，R_3＝H，改变了原为烷基酰胺基取代的类型，为甲氧基酰胺基。由方程（8.18）计算 $\log 1/C = 6.843$。给化学合成，药理试验测得 $\log 1/C = 6.431$，相差 0.412，此值在方程(8.18)的标准偏差 s 范围内，表明预测的结果还是相当准确的。

据文献报导[101,102]，在预测同源物的活性方面有许多成功的例子，读者可参见有关综述[103]。

(2) 有目的地提高化合物的选择性作用。提高化合物的生物活性，同时降低其毒副作用，对药物分子的设计至关重要。应用 Hansch 途径，对一系列同源物的主作用活性与毒副作用的定量数据，往往可以导得不同取代基参数拟合的相关方程，根据不同关系式提示的信息，即可用于定向合成，从而达到提高化合物选择性作用的目的。这种研究又称定量构-择关系研究（quantitative structure-selectivity relationships，简称 QSSR）.已应用于实际工作中。例如，亚硝基脲类抗白血病L-1210 的活性以及毒副作用的相关方程分别为[104]



对白血病细胞 L-1210 的抑制活性：

$$\log 1/C = 1.78 - 0.13(\pm 0.07)\log P - 0.014(\pm 0.015)(\log P)^2 - 0.76(\pm 0.15)I_1 + 0.33(\pm 0.17)I_2 - 0.24(\pm 0.11)I_3 \qquad (8.19)$$

$$n = 90;\ r = 0.868;\ s = 0.206;\ \log P_0 = -4.4$$

毒副作用：

$$\log 1/C = -1.01 - 0.041(\pm 0.007)(\log P)^2 - 0.62(\pm 0.15)I_1 \qquad (8.20)$$

$$n = 96;\ r = 0.829;\ s = 0.221;\ \log P_0 = 0$$

式中，I_1、I_2、I_3 均为指示变量。(8.19)、(8.20) 二式表明其主活性与毒副作用的主要差别在于 $\log P_0$ 不同。抑制 L-1210 白血病细胞的最适宜 $\log P$ 为 −4.4，而毒副作用的 $\log P_0$ 为零。因而设计亲水性强的化合物将具有较高的选择性。

(3) 有助于研究药物的药理作用机制。药物分子产生药理作用，其结构的基本骨架必须与酶或受体的要求相适应。取代基对于化合物与酶或受体的结合将带来显著的影响，而取代基的影响又集中地体现在疏水性、电性和立体效应的变化。QSAR研究中的 Hansch 途径认为理化参数可以模拟药物分子与酶或受体相结合的作用。因此从某种意义上讲，QSAR 导出的最佳方程就代表该类化合物与酶或受体结合的方式。这对于了解药物的作用机制具有很重要的参考价

值。根据 QSAR 研究所得到的方程还可以帮助推测酶或受体图象，并配以量子药理学构象计算和分子图形学的技术，对研究药物作用机制将发挥愈来愈显著的作用。

8.6.5 其他药物分子设计方法

药物设计中，除了 Hansch 途径外还有许多其他方法，它们亦均归类于定量药物设计的范畴之内。主要有：

(1) Topliss方法[105]。Topliss 根据 Hansch 途径中超热力学的概念提出了类似物合成的先后次序的经验方案，旨在合成最大活性的类似物。他提出的合成树方案，在进一步设计需要合成的类似物之前，需要合成和试验(前步)每个化合物，将生物试验的结果与到目前为止所合成的化合物的理化性质相比较，观察哪些理化性质对活性影响起关键作用，(如常选用取代基疏水常数 π)，应用这种判断设计下面一类化合物。合成和试验后，再比较理化性质与生物活性，用以设计再下一个化合物。

Topliss 法对先导化合物进行结构改造阶段有指导作用，它能在合成初期提高找到活性最大的化合物的机会。同时它也为下一步用 Hansch 途径推导有指导意义的QSAR方程打下了基础。近年来 Austel[106] 也发表了与 Topliss 相同原理的所谓系统药物设计的手工方法。

(2) Free-Wilson 途径[107]。QSAR 研究中除了 Hansch 途径外，应用最广的方法是由 Free 和 Wilson 提出经藤田稔夫改进的 De Novo 模型，称 Free-Wilson 途径。

Free-Wilson 途径认为，化合物的生物活性为母体化合物的贡献和各取代基贡献的总和；而以某特定位置上某个基团的有无作为参数。

$$\log 1/C = \mu + \sum^{i}\sum^{j} X_{ij} \cdot G_{ij} \tag{8.21}$$

式中 μ 为母体化合物的生物活性，X_{ij} 系指 ij 取代基存在与否，分别以 1,0 表示，G_{ij} 为基团对生物活性的贡献。它是依据线性回归的方法通过特设的计算机程序来建立的。

(3) 分子连接性法(molecular connectivity)[108]。1975 年Kier 提出了分子连接性法，在以后的药物设计中也获得了广泛的应用。

分子连接性法在计算各个化合物的子图(subgraph)的分子连接性指数 χ(chi) 的基础上，以该指数作为参数来建立 QSAR 方程式，同样具有预测价值。它的显著特点是表达式简单，参数(即分子连接性指数)不是实验值，而是通过代数运算得到。但该参数缺乏较明确的物理意义，使其应用受到限制，如果将该法与 Hansch 途径结合起来使用，可望获得有指导意义的 QSAR 结果。

(4) 模式识别法(pattern recognition)[109]。Kowalski 首先将模式识别法应用到药物设计之中，所谓模式识别法就是借助于电子计算机探索药物大量多元数据或信息中前所未知的相互关系(模式)。它包括研究活性内在规律；对大量观测数据的恒定性质进行检测和识别，为新理论新假说提供依据；对结构相异的化合物提取不同药理特征所赖以的结构信息，以推导其构效关系方程等。

模式识别法的内容十分丰富，它同时综合了药物设计中的其他多种多元统计方法，如因子分

析法，主成分分析法，聚类分析法和判别分析法等内容，通过计算机对约定的化合物进行模式识别，对待测的化合物进行模式分类，这方面内容已有不少专著介绍。

随着计算机技术的进步和普及，QSAR 研究正与量子药理学、X 衍射结晶或分子力学计算机图形学的研究互相配合渗透。新近 Fujita 指出，QSAR 研究已经成功地找到临床应用的新药。它们是用于抗菌的氟哌酸药物——Norfloxacin，抗忧郁药 RS-2232，血管扩张药 Benzyldiphenylmethylpiperazine，抗高血压药 MY-5561[110]。这无疑是十分令人鼓舞的。

正如任何新兴理论一样，在发展中难免存在有待解决的问题。QSAR 研究不能替代药物设计的所有工作，当前还仍大量地进行回顾性研究，预测生物活性并非都能获得成功，更不能靠它发现结构全新的先导化合物，Hansch 途径中所采用的理化参数也并非药物与受体的真正作用力，而是一种模拟，各种方法的物理意义亦有待进一步开发和深化。然而，上述问题必将在今后的发展中得到解决[111]。预料 QSAR 作为药物分子设计的一种理论和实验手段必将得到更广泛的应用和迅速发展。

参 考 文 献

[1] Richards, W. G.: "Quantum Pharmacology" Butterworth & Co. Ltd., 1977.
卓济苍，苏怀德译，科学出版社，(1981).

[2] 嵇汝运：化学通报，4:19 (1984).

[3] Korolkovas, A 著，上海药物研究所《分子药理学概论》翻译小组译："分子药理学概论"，科学出版社，(1976).

[4] 嵇汝运，吴吉安：金荫昌主编：《分子药理学》，天津科技出版社，(1988).

[5] 嵇汝运：邹岗主编《神经药理学》，人民卫生出版社，(1988).

[6] Christoffersen, R. E.: "Quantum Mechanics of Molecular Conformations". Ed by Pullman, B. John Wiley & Sons Ltd. London, P193 (1976).

[7] Richard, W. G.: "Quantum Pharmacology", Second edition, Butterworth & Co. Ltd. Lodon (1983).

[8] Aslanian, D. et al: Mol. Pharm., 22:678 (1982).

[9] 吴加金：军事医学科学院院刊，10:345 (1986).

[10] Epifanir, E. et al: J. Med. Chem., 26:254 (1983).

[11] 嵇汝运：自然杂志，6:685 (1983).

[12] Kelly, J. A. et al: Science, 218:479 (1982).

[13] Cohen, N. C.: J. Med. Chem., 26:259 (1983).

[14] Burkert, U., Allinger, N. L.: "Molecular Mechanics", ACS monograph 177, American Chemical Society Washington, D. C. (1982).

[15] Giglio, E.: Nature, 222:339 (1969).

[16] Gresh, N., Pullman, A.: Theor. Chim. Acta, 66:1 (1984).

[17] Gresh, N., Pullman, A.: Int. J. Quantum Chem., 29:101 (1986).

[18] 吴吉安，陈凯先，嵇汝运：分子科学与化学研究，4:459 (1984).

[19] Pullman, B. et al: Theor. Chim. Acta, 36:317 (1975).

[20] Richards, W. G.: "Steric Effects in Biomolecules", ed Naray-Szabo, G. P61 (1981).

[21] Richards, W. G. et al: J. Theoret. Biol., 51:237 (1975).

[22] Farnell, L. et al: J. Med. Chem., 18:662 (1975).

[23] De, A. U. et al: Indian J. Chem., 20B:58 (1981).

[24] Grol, C. J. et al: J. Pharm. Pharmac., 29:153 (1977).

[25] Harold, L. et al: Proc. Natl. Acad. Sci. USA, 75:2641 (1978).

[26] Costall, B. et al: J. Pharm. Pharmacol., 34:246 (1982).

[27] Philipp, A. H. et al: J. Med. Chem., 22:768 (1979).
[28] Tollenaere, J. P. et al: "Drug Design", Ariens, E. J. Ed., Academic, New York, Vol 10:71 (1980).
[29] Berman, H. M. et al: Proc. Natl. Acad. Sci. USA, 75:828 (1978).
[30] Keepers, J. W. et al: Biopolymers, 23:2901 (1984).
[31] Chen, K. -X, Gresh, N., Pullman, B.: J. Biomol. Struct. Dynam.,. 3:445 (1985).
[32] Chen, K. -X., Gresh, N., Pullman, B.: Mol. Pharmaco., 30:279 (1986).
[33] Chen, K. -X., Gresh, N., Pullman, B.: Nucleic. Acids. Res., 14:2251 (1986).
[34] Chen, K. -X., Gresh, N., Pullman, B.: Nucleic.Acids Res., 14:3799 (1986).
[35] Chen. K. -X., Gresh, N., Pullman, B.: Nucleic. Acids Res., 14:9103 (1986).
[36] Chen. K. -X., Gresh, N., Pullman, B.: Biopolymers, 26:831 (1987).
[37] Chen. K. -X., Gresh, N., Pullman, B.: Nucleic. Acids Res., 15:2175 (1987).
[38] Langlet, G.: J. Appl. Crystallogr, 5:66 (1972).
[39] Andrews, P. R. et al: J. Pharm. Pharmacol., 35:516 (1983).
[40] Andrews, P. R. et al: Med. Res. Rev., 2:355 (1982).
[41] Andrews, P. R. et al: J. Med. Chem., 29:453 (1986).
[42] Andrews, P. R. et al: Quant. Struct.-Act. Relat., 6:97 (1987).
[43] Duvernay, V. H. et al: Mol. Pharmacol., 16:623 (1979).
[44] 嵇汝运：国外医学药学分册，(5):257 (1983).
[45] 顾坤健：《量子化学译文集》，上海科技文献出版社 P45 (1981).
[46] Kang, S. et al: Mol. Pharmacol., 9:640 (1973).
[47] Mazurek, A. P.; Weinstein, H. et al: Int. J. Quantum Chem. Quantum Biol. Symp., 11:183 (1984).
[48] Curtiss, L. et al: Chem. Phys, Lett., 103:437 (1984).
[49] Matthews, D. A. et al: Science, 197:452 (1977).
[50] Kuyper, L. F. et al: J. Med. Chem., 25:1120 (1982).
[51] Burchall, J. J.: J. Infeet. Dis., 128 (Nov) (Suppl):S437 (1973).
[52] Hansch, C. et al: J. Med. Chem., 20:96 (1977).
[53] Hansch, C. et al: J. Med. Chem., 25:777 (1982).
[54] Li R, -L. et al: Quant. Struct. -Act. Relat., 1:1 (1982).
[55] Arup, K. et al: J. Med. Chem., 28:333 (1985).
[56] Andrews, P. R. et al: J. Med. Chem., 29:698 (1986).
[57] Millex, K. J. et al: QCPE Bull., 1:77 (1981).
[58] Matthews, D. A. et al: J. Biol. Chem., 260:381 (1985).
[59] Giessner-Prettre, C.; Pullman, A.: Theor. Chim. Acta, 25:83 (1982).
[60] Czaplinsky, K. -H. et al: Quant. Struct. -Act. Relat., 6:70 (1987).
[61] Gold. E. H. et al: J. Med. Chem., 25:1363 (1982).
[62] McGrath, J. C.: Biochem. Pharmacol, 31:467 (1982).
[63] Carpy, P. et al: Mol. Pharmacol., 21:400 (1982).
[64] 吴吉安，陈凯先，嵇汝运：分子科学学报，2:27 (1981).
[65] 郭直惟，李正化：华西医药杂志，1:74 (1986).
[66] Bohl, M. et al: Eur. J. Med. Chem., 21:193 (1986).
[67] Bennett, L. K. et al: J. Pharm. Sci., 75:769 (1986).
[68] Liebman, M. N.: J. Mol. Graphics, 4:61 (1986).
[69] 陈凯先，吴吉安，嵇汝运：中国科学（B），12:1257 (1986).
[70] Durant, G. J. et al: J. Med. Chem., 18:905 (1975).
[71] 陈凯先，嵇汝运等：分子科学与化学研究，1:25 (1983).
[72] Petrongolo, C. et al: J. Med. Chem., 20:1645 (1977).
[73] Macchia, B. et al: Eur. J. Med. Chem., 15:515 (1980).
[74] Adams, G. E. et al: Int. J. Radiat. Biol., 15:457 (1969).

[75] Adams, G. E. et al: Radiat. Res., 9:67 (1976).
[76] Lin, L. -M.; Xu, S. -J.; Lu, Z. -Y.; Wu, J. -A.: Radiosensitization Newsletter, 3:2 (1984).
[77] Willson, R. L. et al: Int. J. Radiat. Biol., 26:557 (1974).
[78] Lin, L. -M.; Xu, S. -J.; Lu, Z. -Y.; Wu, J. -A.: Radiosensitization Newsletter, 4:2 (1985).
[79] 刘泰福等：待发表
[80] Xu. S. -J.; Li, L. -M.; Lu, Z. -Y.; Wu, J. -A.: J. Mol Sci., (Wuhan China) 4:11 (1986).
[81] Inayma, S.: Proc. of the 2nd Japan-China Bilateral Symp. on Radiosensitization, 84 (1986).
[82] van de Waterbeemd, H. et al: J. Med. Chem., 28:203 (1983).
[83] Kikuchi, O. et al: Cancer Biochem. and Biophys. 4:1 (1979).
[84] 白迺彬，曲绍清：分子科学学报，2:97 (1982).
[85] 永田亲义著，陶宗晋，江寿平译：《量子生物学入门》，上海科技出版社 P 182 (1979).
[86] Sabin, K. et al: J. Pharm. Sci., 69:796 (1980).
[87] Cheney, B. V. et al: Int. J. Quantum Chem. Quantum Biol. Symp., 8:201 (1981).
[88] Hansch, C. and Fujita, T.: J. Am. Chem. Soc., 86:1616 (1964)., Hansch, C.: Adv. Pharm. Chemother., 13:45 (1975). Hansch, C.: J. Med. Chem., 19:1 (1976).
[89] Hansch, C.: "Drug Design", Ariens, E. J. Ed., Vol 1, Academic, New York P271, (1971).
[90] 藤田稔夫等：化学の领域增刊 122 号，(1979).
藤田稔夫等：化学の领域增刊 136 号，(1982).
[91] 王尔华编译：《定量药物设计》，人民卫生出版社 (1983).
[92] 吴吉安，嵇汝运：中国药理学报，3:51 (1982).
[93] Kubinyi, H. et al: Arzeim-Forsch, 28:598 (1978).
Kubinyi, H. et al: Farmaco Ed. Sci., 34:248 (1979).
[94] 吴吉安，嵇汝运：计算机与应用化学，1:273 (1984).
[95] Hansch, C. and Leo, J.: "Substituent Constants for Correlation Analysis in Chemistry and Biology", J. Wiley and Sons, New York (1979).
[96] Rekker, R. F.: "The Hydrophobic Fragmental Constant in Biological Activity and Chemical Structure", Buisman, J. A. K. Ed. Elsevier, Amsterdam (1978).
[97] Leo, A. Pomona College Medicinal Chemistry Project Pomona, Calif., (1981).
[98] 张仁斌，徐修容：《高效液相色谱在医药研究中的应用》，上海科技出版社 (1983).
[99] 吴吉安，顾坤健，嵇汝运：药学学报，16:828 (1981).
[100] 朱友成，吴吉安，徐修容：药学学报，20:267 (1985).
[101] Li, R. -L. et al: J. Med. Chem., 24:538 (1981).
[102] Martin, Y. C. et al: J. Med. Chem., 24:229 (1981).
[103] 李仁利：国外医学药学分册. (2):65 (1983).
[104] Hansch, C. et al: J. Med. Chem., 23:1095 (1980).
[105] Topliss, J. G.: J. Med. Chem., 15:1006 (1972).
[106] Austel, V.: Eur. J. Med. Chem., 17:9 and 17:339 (1982).
[107] Purcell, W. P. et al: "Strategy of Drug Design", Wileg, New York (1973).
[108] Kier, L. B. et al: "Molecular Connectivity in Chemistry and Drug Research", Academic Press, New York (1976).
[109] Kowalski, B. r. et al: J. Amer. Chem. Soc., 94:5632 (1972).
[110] Fujita, T.: "10 th International Congress of Pharmacology", Sydney Australia S43 (1987).
[111] Seydel, J. K.: "QSAR and Strategies in the Design of Bioactive Compounds", VCH Verlagsgesellschaft mbH, D-6940 Weinhein (1985).

吴吉安

第9章　生物分子的空间构象

生物分子的空间构象问题，在第3、第4和第5章里，结合分子间相互作用，蛋白质和核酸的空间构象对其生物学功能影响的讨论，已经作了初步的介绍。但是，鉴于生物分子（尤其是生物大分子）的空间构象与功能直接相关，特别是从研究生物分子空间构象本身出发，来探讨在溶液中对构象的变化，生物分子的折叠规律以及在此基础上的构象模拟和预测，更具重大的意义。为此，我们专辟一章来比较详尽地讨论生物分子的空间构象问题[1-4]。

回顾有关空间构象的历史，我们可以清楚地看到，首先是X射线晶体衍射的研究，为阐明生物分子的三维结构奠定了基础。但是，X射线衍射分析是相当费时、费力的，而且尚有许多重要生物分子迄今还未能得到合适的结晶。其它一些技术，如旋光色散和圆二色谱虽然可以得到溶液状态蛋白质分子二级结构的信息（如α螺旋和β折叠结构的含量），但它既不能提供这些结构的起止残基位置（只能提供一个百分数），更不能提供三维结构的信息。近些年发展起来的二维NMR谱，虽为研究溶液中生物分子的空间构象提供了一个新的方法，但目前还只能对小的蛋白质分子进行结构分析，且也相当费时。从另一方面来看，如果能够发展和运用理论计算方法，那也将是一条不可缺少的途径。况且随着近代物理学、数学、化学和分子生物学的发展，特别是计算机技术的发展，遗传工程、蛋白质工程和生物医学工程的发展，迫切要求发展理论计算方法，并已初步显示出这样的方法对模拟和预测生物分子空间构象具有独特的优点以及与实验测定比较接近的结果。

9.1　蛋白质和核酸分子二级结构的预测

9.1.1　蛋白质分子二级结构的预测

从蛋白质的一级结构来预测它的二级结构有许多不同的方法。如Scheroga和他的合作者[5]基于能量计算方法，预测了4个蛋白质的α螺旋和无规卷曲，准确率达到78%。Robson及其同事[6]应用信息论方法，而Lim[7]则基于一系列立体化学的规律分别对蛋白质分子的二级结构进行了研究。但是相对比较简单，应用较广的还是Chou-Fasman[8]的预测方法，因而以下我们将着重介绍他们的方法。

1. α螺旋

蛋白质链构象可用二面角ϕ和ψ来描述，对于一个正规的右手α螺旋残基的ϕ、ψ角是$(-57^\circ,-47^\circ)$。α_{I}和α_{II}螺旋构象的ϕ和ψ各是$(-49^\circ,-26^\circ)$和$(-93^\circ,-18^\circ)$常出现在螺旋末端。一些残基螺旋末端要是含有氢键或许二面角在$-100^\circ\leqslant\phi\leqslant-20^\circ$和$-90^\circ\leqslant\psi\leqslant10^\circ$范围之外。图9-1（$A$）是对一些蛋白质所含的螺旋统计的结果，4个螺旋有4个残基，3个最长的螺旋含有26个残基。有24个螺旋片段的残基数长于17个残基。

2. β 折叠区域

平行和反平行 β 折叠残基的 ϕ 和 ψ 值各是(−119°,113°)和(−139°,135°)。连续 3 个或更多残基序列有 ϕ 和 ψ 各在(−120°,110°)或(−140°,135°)的 40°之内认为是处于 β 构象。然而,在 β 折叠的末端残基,有两面角在 $-180°\leqslant\phi\leqslant+180°$ 和 $175°\leqslant\psi\leqslant70°$ 范围之外,并至少有一个氢键也包括在 β 区域。一些蛋白质的 β 区域表示在图 9-1(*B*)中, 有 3 个只有 2 个残基的 β片段。

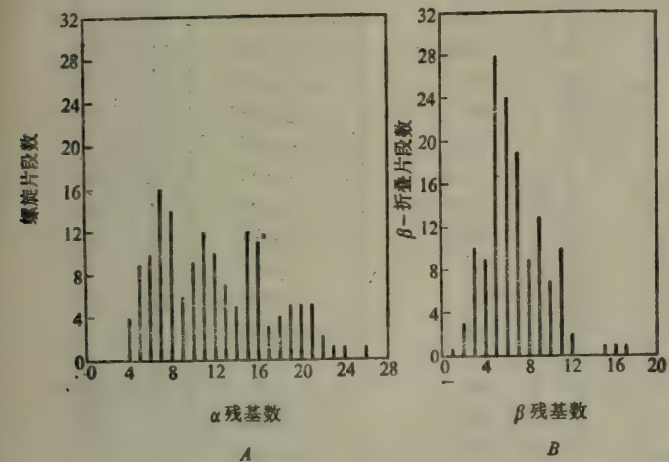


图 9-1 每螺旋片段(*A*)和 β 折叠片段(*B*)的残基数分布

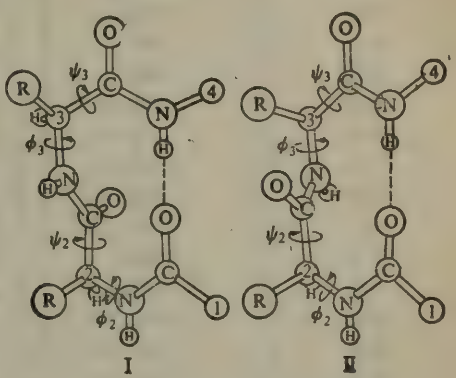


图 9-2 转角结构

3. β 转角区域

在蛋白质中,β 转角包含着 4 个残基,其多肽链本身几乎以180°折回。正是由于这些链的反转区域,使蛋白质分子形成球状,而不是线形的分子。3 种以四肽形成的转角,它们的残基 i 的 CO 基团和残基 $i+3$ 的 NH 基团之间形成氢键,图 9-2 表示了二种型式的转角。型式 I 转角的 $(\phi,\psi)_2=(-60°,-30°)$和$(\phi,\psi)_3=(-90°,-0°)$,而型式 II 转角的$(\phi,\psi)_2=(-60°,120°)$和$(\phi,\psi)_3=(80°,0°)$,下指标 2 和 3 各指 β 转角的第二和第三残基。$O_1—N_4$的距离小于 0.35 nm 和 $C_1^{\alpha}—C_4^{\alpha}$距离小于 0.7 nm。型式 I 绕第二肽键转动180°则为型式 II。型式转角的第二和第三残基的两面角一样,它们的$(\phi,\psi)=(-60°,-30°)$。

9.1.1.1 Chou 和 Fasman 方法[8]

Chou 和 Fasman 对应用 X 光衍射得到的 29 个蛋白质数据进行了统计从而得到各种氨基酸残基在蛋白质中形成 α 螺旋,β 折叠的构象参数。表 9-1 为各个氨基酸形成 α 螺旋和β 折叠的构象参数,表 9-1 中以它们形成 α 螺旋或 β 折叠的趋势大小顺序排列。

表中 P_α和 P_β各为 α 螺旋和 β 折叠的构象参数,H_α和 H_β各是 α 和 β 强的形成者;h_α和h_β各是 α 和 β 形成者;I_α和 I_β各为 α 和 β 弱的形成者;i_α和 i_β为无形成 α 和 β 者;b_α和 b_β各为 α 和 β 破坏者;B_α和 B_β各为 α 和 β 强破坏者。此表表示了最易形成 α 螺旋的残基是谷氨酸、蛋氨酸、丙氨酸和亮氨酸,而缬氨酸、异亮氨酸和酪氨酸是最易形成 β 折叠的残基。甘氨酸、脯氨酸、天冬酰胺和酪氨酸是破坏 α 螺旋的残基。而谷氨酸是最易破坏 β 折叠的残基。5 个带电

表 9-1 α螺旋和β折叠的构象参数

α-残基	P_α	螺旋分配	β-残基	P_β	片层分配
Glu^-	1.51	H_α	Val	1.70	H_β
Met	1.45	H_α	Ile	1.60	H_β
Ala	1.42	H_α	Tyr	1.47	H_β
Leu	1.21	H_α	Phe	1.38	h_β
Lys^+	1.16	h_α	Trp	1.37	h_β
Phe	1.13	h_α	Leu	1.30	h_β
Gln	1.11	h_α	Cys	1.19	h_β
Trp	1.08	h_α	Thr	1.19	h_β
Ile	1.08	h_α	Gln	1.10	h_β
Val	1.06	h_α	Met	1.05	h_β
Asp^-	1.01	I_α	Arg^+	0.93	i_β
His^+	1.00	I_α	Asn	0.89	i_β
Arg^+	0.98	i_α	His^+	0.87	i_β
Thr	0.83	i_α	Ala	0.83	i_β
Ser	0.77	i_α	Ser	0.75	b_β
Cys	0.70	i_α	Gly	0.75	b_β
Tyr	0.69	b_α	Lys^+	0.74	b_β
Asn	0.67	b_α	Pro	0.55	B_β
Pro	0.57	B_α	Asp^-	0.54	B_β
Gly	0.57	B_α	Glu^-	0.37	B_β

荷的残基都不利于β折叠的形成。

虽然表 9-1 中的 P_α和 P_β表明了在蛋白质中形成α螺旋和β折叠的相对可能性，但并没有指出这些残基在它们各自结构区域中位置的择优。通过构象边界分析，可以得到哪一个残基更常出现在α螺旋和β折叠的 N 端和 C 端。表 9-2 和 9-3 是通过对 29 个蛋白质分析得到的α螺旋和β折叠边界残基的构象参数。表中 $P_{\alpha N}$和$P_{\alpha C}$各为在螺旋区域 N 端和 C 端的残基归一化频率，$P_{n\alpha N}$和 $P_{n\alpha C}$各为在非螺旋区域 N 端和 C 端残基的归一化频率。$P_{\beta N}$，$P_{\beta C}$，$P_{n\beta N}$ 和 $P_{n\beta C}$类推之。从表 9-3 可以看出，负电荷残基谷氨酸和天冬氨酸出现在 N 端螺旋末端占优势，而正电荷残基赖氨酸、组氨酸和精氨酸出现在 C 端螺旋末端占优势。脯氨酸残基从不出现在螺旋 C 端，而常出现在螺旋的 N 端。表 9-3 还表明，带电残基不常出现在β折叠的边界区域，除赖氨酸以外，所有其它带电残基的 $P_{\beta N}<1.00$ 和 $P_{\beta C}<1.00$。谷氨酸、天冬氨酸和组氨酸的 $P_{\beta N}$值最小。异亮氨酸、缬氨酸、谷氨酰胺和苯丙氨酸易出现在β折叠 N 端，而酪氨酸、缬氨酸和苯丙氨酸等则易出现在β折叠的 C 端。

Chou 和 Fasman 同样基于对已知 X 光结构的 29 个蛋白质的分析表明，在链的反转区域有 408 个β转角(或称回折，发夹结构等)，β转角的第一到第四个残基的频率按其大小列于表 9-4 中。表中 i，$i+1$，$i+2$，$i+3$ 各代表在一个反转β转角处第一，二，三和四个残基的频率。P_t是基于转角所有四个位置上的一个残基在β转角中的构象势。P_{t2}是基于一个反转转角的第二和第三位置上的一个残基在β转角中的构象势。表 9-4 表明，出现在 i 位置频率最大的有天冬酰

表 9-2　螺旋边界残基的构象参数

	$P_{\alpha N}$	$P_{\alpha C}$	$P_{n\alpha N}$	$P_{n\alpha C}$
Glu$^{(-)}$	2.44 Lys$^{(+)}$	1.83 Ser	1.55 His$^{(+)}$	1.86
Asp$^{(-)}$	2.02 His$^{(+)}$	1.77 Asn	1.42 Asn	1.64
Pro	2.01 Met	1.57 Gly	1.41 Gly	1.64
Trp	1.47 Val	1.25 His$^{(+)}$	1.22 Pro	1.58
Ala	1.29 Arg$^{(+)}$	1.25 Pro	1.10 Lys$^{(+)}$	1.49
Gln	1.22 Glu$^{(-)}$	1.24 Thr	1.09 Arg$^{(+)}$	1.24
Thr	1.08 Gln	1.22 Glu$^{(-)}$	1.04 Asp$^{(-)}$	1.06
Asn	0.81 Ala	1.20 Lys$^{(+)}$	1.01 Phe	1.04
Gly	0.76 Leu	1.13 Tyr	0.99 Tyr	0.96
Ser	0.74 Cys	1.11 Asp$^{(-)}$	0.98 Cys	0.94
His$^{(+)}$	0.73 Phe	1.10 Phe	0.93 Ser	0.93
Met	0.71 Ile	0.98 Leu	0.85 Ile	0.87
Tyr	0.68 Ser	0.96 Met	0.83 Thr	0.86
Ile	0.67 Thr	0.75 Ile	0.78 Leu	0.84
Cys	0.66 Tyr	0.73 Gln	0.75 Gln	0.70
Lys$^{(+)}$	0.66 Asp$^{(-)}$	0.61 Val	0.75Glu$^{(-)}$	0.59
Phe	0.61 Asn	0.59 Ala	0.70 Ala	0.52
Val	0.61 Gly	0.42 Cys	0.65 Met	0.52
Leu	0.58 Trp	0.40 Trp	0.62 Val	0.32
Arg$^{(+)}$	0.44 Pro	0.00 Arg$^{(+)}$	0.34 Trp	0.16

表 9-3　β折叠边界残基的构象参数

	$P_{\beta N}$	$P_{\beta C}$	$P_{n\beta N}$	$P_{n\beta C}$
Ile	1.94 Tyr	1.96 Asn	1.86 Pro	1.69
Val	1.69 Val	1.79 Pro	1.58 Gly	1.68
Gln	1.65 Phe	1.50 Gly	1.46 Trp	1.59
Phe	1.50 Ile	1.35 Ser	1.41 Ser	1.49
Trp	1.49 Leu	1.27 Asp$^{(-)}$	1.39 Asp$^{(-)}$	1.32
Met	1.43 Asn	1.21 Cys	1.34 Thr	1.16
Leu	1.30 Trp	1.19 Tyr	1.23 Asn	1.13
Thr	1.17 Cys	1.11 Lys$^{(+)}$	1.09 Arg$^{(+)}$	1.05
Tyr	1.07 Met	0.95 Gln	1.09 Tyr	1.01
Lys$^{(+)}$	1.00 His$^{(+)}$	0.90 Thr	1.09 His$^{(+)}$	0.96
Arg$^{(+)}$	0.90 Arg$^{(+)}$	0.90 Clu$^{(-)}$	0.92 Met	0.85
Cys	0.87 Asp$^{(-)}$	0.85 Arg$^{(+)}$	0.89 Glu$^{(-)}$	0.85
Ala	0.86 Ser	0.79 His$^{(+)}$	0.78 Lys$^{(+)}$	0.82
Pro	0.66 Thr	0.75 Ala	0.67 Gln	0.77
Asn	0.66 Ala	0.75 Ile	0.59 Ala	0.74
Gly	0.63 Gly	0.74 Met	0.52 Val	0.59
Ser	0.63 Lys$^{(+)}$	0.74 Trp	0.48 Leu	0.59
His$^{(+)}$	0.54 Gln	0.65 Leu	0.46 Ile	0.53
Asp$^{(-)}$	0.38 Glu$^{(-)}$	0.55 Val	0.42 Cys	0.53
Glu$^{(-)}$	0.35 Pro	0.40 Phe	0.30 Phe	0.44

表 9-4 β转角中氨基酸出现频率

	i	i+1	i+2	i+3	P_t	P_{te}
Asn	0.161 Pro	0.301 Asn	0.191 Trp	0.167 Asn	1.56 Pro	2.04
Cys	0.149 Ser	0.139 Gly	0.190 Gly	0.152 Gly	1.56 Gly	1.63
Asp	0.147 Lys	0.115 Asp	0.179 Gys	0.128 Pro	1.52 Asp	1.61
His	0.140 Asp	0.110 Ser	0.125 Tyr	0.125 Asp	1.46 Asn	1.56
Ser	0.120 Thr	0.108 Cys	0.117 Ser	0.106 Ser	1.43 Ser	1.52
Pro	0.102 Arg	0.106 Tyr	0.114 Gln	0.098 Cys	1.19 Lys	1.13
Gly	0.102 Gln	0.098 Arg	0.099 Lys	0.095 Tyr	1.14 Tyr	1.08
Thr	0.086 Gly	0.085 His	0.093 Asn	0.091 Lys	1.01 Arg	1.05
Tyr	0.082 Asn	0.083 Glu	0.077 Arg	0.085 Gln	0.98 Thr	0.98
Trp	0.077 Met	0.082 Lys	0.072 Asp	0.081 Thr	0.96 Cys	0.92
Gln	0.074 Ala	0.076 Thr	0.065 Thr	0.079 Trp	0.96 Gln	0.84
Arg	0.070 Tyr	0.065 Phe	0.065 Leu	0.070 Arg	0.95 Glu	0.80
Met	0.068 Glu	0.060 Trp	0.064 Pro	0.068 His	0.95 His	0.77
Val	0.062 Cys	0.053 Gln	0.037 Phe	0.065 Glu	0.74 Ala	0.64
Leu	0.061 Val	0.048 Leu	0.036 Glu	0.064 Ala	0.66 Phe	0.62
Ala	0.060 His	0.047 Ala	0.035 Ala	0.058 Met	0.60 Met	0.51
Phe	0.059 Phe	0.041 Pro	0.034 Ile	0.056 Phe	0.60 Trp	0.48
Glu	0.056 Ile	0.034 Val	0.028 Met	0.055 Leu	0.59 Val	0.43
Lys	0.055 Leu	0.025 Met	0.014 His	0.054 Val	0.50 Leu	0.36
Ile	0.043 Trp	0.013 Ile	0.013 Val	0.053 Ile	0.47 Ile	0.29

胺，半胱氨酸、天冬氨酸；出现在 $i+1$ 位置频率最大的是脯氨酸、丝氨酸、赖氨酸；在 $i+2$ 位置频率最大的为天冬酰胺、甘氨酸和天冬氨酸；而在 $i+3$ 位置的是色氨酸、甘氨酸和半胱氨酸。

某些残基在 β 转角中有戏剧性的位置择优，如脯氨酸出现在第二位置而不出现在第三位置上；色氨酸出现在第四位置上但不在第二位置上；组氨酸出现在第一位置，但不出现在第二或第四位置上；赖氨酸出现在第二但不出现在第一位置上等。在四个位置都有高的 β 转角构象势的有：天冬酰胺、甘氨酸、脯氨酸和丝氨酸。疏水性最大的残基表现出 β 转角构象势最低。而在另一方面接近 β 转角区域却发现许多疏水残基。事实上 β 转角是作为螺旋和 β 折叠的终止者。

有了以上的数据，就可以来确定多肽和蛋白质的二级结构，寻找它们的 α 螺旋和 β 折叠等区域。

9.1.1.2 各种预测方法结果的比较

X 射线衍射在结果发表以前，一些研究组[9]用不同方法预测腺苷酸激酶的二级结构，这些结果表示于图 9-3 中。从图 9-3 可见，螺旋预测结果和实验结果很符合。在这种激酶中总共有 10 个螺旋，Chou 和 Fasman，Fonkelstin 和 Ptitsyn 确定了 9 个。Barry 和 Friedman 以及 Lim 等找到 8 个螺旋。5 个 β 折叠区域，所有的研究组都正确的确定了三个 β 折叠区域：10—14，90—94和114—118残基。但对这两种结构区域的预测有的区域没预测到，有的是预测了实际不存在的区域。由于各种方法对 β 转角的算法几乎相同，所以各种方法预测的结果也很相似。

比较各种方法对腺苷酸激酶二级结构预测的准确性，以 Ptitsyn 与其同事的方法预测螺旋的结果为最好；Lewis 和他的同事的方法预测 β 转角的结果最好；Lim 方法得到最好的 β 折叠的预测结果，而 Chou 和 Fasman 对上述三种构象区域预测结果都居于第二位。

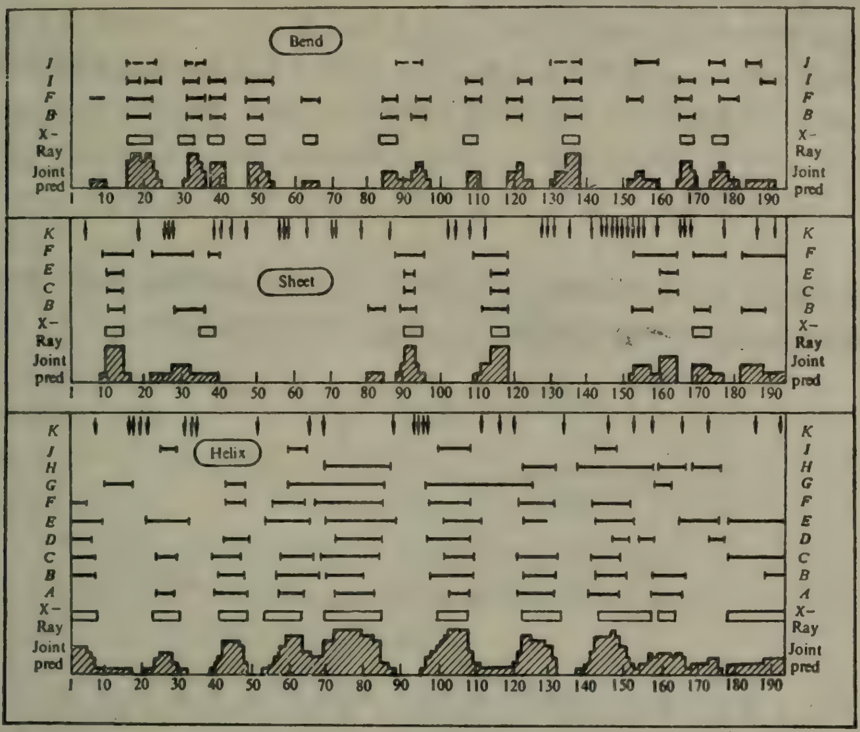


图 9-3 腺苷酸酶二级结构实验结果和各预测方法的比较

J、Burgess-Scheraga，I、Lewis-Scheraga

F、Nagano，B、Chou-Fasman，K、Kabat-w

E、Lim，C、Frnkelstein-Ptitsyn，D、Levi-Robson

H、Kotelchnck-Scheraga，A、Barrs-Friedman

9.1.1.3 蛋白质二级结构预测的计算机程序[10,11]

二级结构的预测，像Chou-Fasman的预测方法是可以手算的，但对于大的蛋白质分子，靠手算既费时又费力，而用计算机则可很快预测出分子的二级结构，而对它进行分析处理，即使对那些不熟悉具体预测方法的人也可以使用，既方便又省时。是实验室中常规方法之一。

我国科学工作者刘志平和江寿平设计的蛋白质二级结构预测的计算机算法，是进行蛋白质三级结构预测的基础。随着有关蛋白质构象方面信息的积累，已发展起来不少预测蛋白质高级结构的理论和方法。由蛋白质的氨基酸顺序进行其二级结构的预测基本上分成三类，第一类是统计方法，它对一组已知X射线分析结果的蛋白质进行采样，计算出 20 种天然氨基酸在α-螺旋，β-折叠，无规结构中出现的频率。并据此赋给每一种氨基酸一构象参数，例如α-螺旋构象势P_α，β-折叠构象势P_β等。然后，再按照一定的规则，用这些参数推测出蛋白质的二级结构。这类方法中，较简便而成功的方法被公认为Chou和Fasman方法。然而，Chou和 Fasman 方法关于处理同

时具有α-螺旋结构势$\langle P_\alpha\rangle$和β-折叠结构势$\langle P_\beta\rangle$的序列所采用的规则 3，在$\langle P_\alpha\rangle$与$\langle P_\beta\rangle$相差不大时显得模棱两可。第二类方法按照立体化学规则直接构造蛋白质的二级结构。近年来，人们发现氨基酸残基的疏水极性对于高级结构的形成起关键作用。因此，有人试图根据序列中残基的疏水极性的大小及极性强弱的周期性来预测蛋白质的二级结构[12]。我们根据立体化学规则及序列中残基的疏水性强弱变化的周期性，提出了解决$\langle P_\alpha\rangle\approx\langle P_\beta\rangle$时的重叠片段的归属的算法，并利用类Chou和Fasman规则 1 及规则 2 提出了适于在计算机上实现预测的计算机算法。此算法以蛋白质的氨基酸顺序或与之对应的核苷酸顺序作为输入，输出蛋白质各二级结构单元在肽链中的位置。由于这一算法的简单、快速且可在微型机上实现，因此，它完全可以作为一种日常分析手段而用于实验室。

1. 理论

在进行蛋白质分子二级结构预测时，我们考虑的不是单个氨基酸残基，而是一段序列。我们可以用 4 种度量来标识这些序列所携带的形成一特定二级结构s所需的信息大小。

$$nf(j)=\sum_{i=j}^{j+n_r} f_s(R_i)\quad j=1,2,\cdots N-n_r\quad s\varepsilon(\alpha,\beta) \tag{9.1}$$

$$nb(j)=\sum_{i=j}^{j+n_k} b_s(R_i)\quad j=1,2,\cdots N-n_r\quad s\varepsilon(\alpha,\beta) \tag{9.2}$$

$$\langle P_s\rangle=\sum_{i=j}^{j+n_r} P_s(R_i)/(n_r+1)\quad j=1,2,\cdots N-n_r\quad s\varepsilon(\alpha,\beta) \tag{9.3}$$

$$H_f(R_i)$$

其中R_i是第i个残基，s表示一特定的二级结构（α-螺旋，β-折叠或无规结构），$nf(j)$表示序列$j-(j+n_r)$所含有的有利于形成s结构的信息。$nb(j)$表示序列$j-(j+n_r)$所含有的不利于形成s结构的信息。$\langle P_s\rangle$表$j-(j+n_r)$中每个残基所携带的形成s结构的平均信息值。$H_f(R_i)$表残基R_i的疏水极性大小。关于上述诸量的取值，定义如下：

$$f_\alpha(R_i)=\begin{cases}1 & P_\alpha(R_i)>1.03\\ 0.5 & 1.0\leqslant P_\alpha(R_i)<1.03\\ 0 & P_\alpha(R_i)\leqslant 1.0\end{cases}$$

$$b_\alpha(R_i)=\begin{cases}1 & P_\alpha(R_i)<0.7\\ 0 & P_\alpha(R_i)\geqslant 0.7\end{cases}$$

$$f_\beta(R_i)=\begin{cases}1 & P_\beta(R_i)>1.0\\ 0 & P_\beta(R_i)\leqslant 1.0\end{cases} \tag{9.4}$$

$$b_\beta(R_i)=\begin{cases}1 & P_\beta(R_i)<0.8\\ 0 & P_\beta(R_i)\geqslant 0.8\end{cases}$$

在蛋白质分子的合成过程中，多肽链的延伸和高级结构的形成是同时进行的。因此，不妨作如下假设：

(1) 在肽链的合成过程中，当合成出来的某一特定顺序$j-(j+n_r)$携带有足够的形成一

定二级结构所需的信息量时，它可优先形成一核。

(2) 随着肽链的延伸，这个核也可向C端延伸。并且由于该核C端后继合成出来的各氨基酸残基与此核N端及N端前面各残基的相互作用，此核也可向肽链N端延伸，当遇到一特定顺序不具有足够的信息量时，可以中断核的延伸，从而定位出蛋白质肽链中各二级结构单元。

在本节所提出的算法中，对于α-螺旋，我们选择了$n_r=5$。取向肽链C端延伸的顺序长度为4肽，向肽链N端延伸的顺序长度为5肽。对于β-折叠，选择$n_r=4$，取向肽链C端延伸的顺序长度为4肽，向肽链N端延伸的顺序长度为5肽。

2. 方法

基于上述理论及假设(1)，(2)，我们设计算法如下：

(1) 输入

此算法为用户提供了二种输入。一为蛋白质的氨基酸顺序。另一为与之对应的核苷酸顺序(连同其读码框架)。当输入为核苷酸顺序时，通过一翻译程序按照三联体密码与氨基酸的对应关系自动翻译成对应的氨基酸顺序。

(2) α-螺旋的寻找

① 从肽链的N端至C端沿肽链寻找满足$nf(j)\geqslant 4$且$nb(j)\leqslant 1$且不含Gly、Pro的6肽片段，并把它定义为α-螺旋核。

② 从核的C端开始，向肽链的C端每4个氨基酸残基一组，逐组计算$\langle P_\alpha\rangle$，$nb(j)$，直至遇到满足$\langle P_\alpha\rangle<1.0$或$nb(j)\geqslant 2$或含有Pro，Gly的4肽顺序，中断延伸，并按照Chou和Fasman的规则1(4)及规则1(5)定出此二级结构单元的C端。

③ 从螺旋核的N端开始，向肽链N端每5个氨基酸残基一组，逐组计算$\langle P_\alpha\rangle$，$nb(j)$，延伸此核，直至遇到满足$\langle P_\alpha\rangle<1.0$或$nb(j)\geqslant 2$，或含有Gly、Pro的5肽顺序。中断延伸，按照Chou和Fasman规则1(4)及规则1(5)定出此二级结构单元的N端。若此片段的$\langle P_\alpha\rangle\gg\langle P_\beta\rangle$且$\langle P_\beta\rangle<1.05$，则此片段可定义为$\alpha$-螺旋结构。

3. β-折叠的寻找

β-折叠的寻找同α-螺旋的寻找类似。作为β-折叠核心的5肽顺序应满足的条件是$nf(j)\geqslant 3$且$nb(j)\leqslant 1$且不含Glu、Pro。中止核向肽链N端延伸的5肽顺序及向C端延伸的4肽顺序应满足的条件$\langle P_\beta\rangle<1.0$或$nb>1.0$或含有Glu、Pro，二级结构单元的N端、C端应满足Chou和Fasman规则2(4)及规则2(5)。若此片段的$\langle P_\alpha\rangle<1.03$且$\langle P_\beta\rangle>\langle P_\alpha\rangle$则可定义为$\beta$-折叠结构。

重复二及三，直到肽链中所有的氨基酸残基都被考察过。

4. 重叠处理

所谓重叠即是由算法二及算法三所预测出来的片段$j_{\alpha N}-j_{\alpha C}$，$j_{\beta N}-j_{\beta C}$发生重叠。重叠部分$j_{N'}-j_{C'}$ ($j_{N'}\geqslant \mathrm{maximum}(j_{\alpha N}, j_{\beta N})$，$j_{C'}\leqslant \mathrm{minimum}(j_{\alpha C}, j_{\beta C})$) 既具有$\alpha$-螺旋势又具有$\beta$-折叠势。为处理重叠区域$j_{N'}-j_{C'}$的最终归属，我们考虑下列几点事实：

(1) α-螺旋和位于球蛋白表面的β-折叠通常具有面向溶剂的极性表面及使它们形成三级结构时相互堆积在一起的非极性表面(由疏水性强的氨基酸残基组成)。而且α-螺旋中的侧链

基团的极性强弱的周期性与β-折叠的不一样。前者显周期为2—3个残基的变化，而后者显周期为一个残基的变化。

(2) 埋入球蛋白内部的β-折叠常常是由若干连续排列在一起的疏水性较强的氨基酸残基组成。因此，我们可以通过考察重叠片段中残基的疏水性强弱及其周期性变化情况来区分重叠片段的最终归属。具体算法如下：

① 当$j_{\beta N}-j_{\beta C}$处在$j_{\alpha N}-j_{\alpha C}$的中间($j_{N'}=j_{\beta N}>j_{\alpha N}, j_{C'}=j_{\beta C}<j_{\alpha C}$)则$j_{\beta N}-j_{\beta C}$片段应为$\alpha$-螺旋结构。因为形成$\alpha$-螺旋结构比形成$\beta$-折叠结构在能量上更为有利。

② 当$j_{\beta N}-j_{\beta C}$不处在$j_{\alpha N}-j_{\alpha C}$的中间，如处在$j_{\alpha N}-j_{\alpha C}$的两末端或与$j_{\alpha N}-j_{\alpha C}$发生交叉重叠且重叠片段的长度大于3个残基或$j_{\beta N}-j_{\beta C}$包含$j_{\alpha N}-j_{\alpha C}$。分析重叠部分的疏水极性强弱的周期性，若重叠部分有连续排列的其疏水极性值$H_r(R_i)>12.0$Kcal的4个以上氨基酸残基，或其氨基酸残基的疏水极性值显$H_r>12.0$，$H_r\leqslant 12$(Kcal)的交替变化，则此重叠片段为β-折叠结构[12]。否则，为α-螺旋结构。

③ 若$j_{\beta N}-j_{\beta C}$与$j_{\alpha N}-j_{\alpha C}$发生交叉重叠且重叠部分的长度小于4个氨基酸残基，则利用20种氨基酸的边界频率——出现在N末端非α-螺旋区的频率$f_{n\alpha N}$，非β-折叠区的频率$f_{n\beta N}$，出现在N末端α-螺旋区的频率$f_{\alpha N}$，β-折叠区的频率$f_{\beta N}$，出现在C末端非α-螺旋区，非β-折叠区的频率$f_{n\alpha C}$及$f_{n\beta C}$，出现在C末端α-螺旋区，β-折叠区的频率$f_{\alpha C}$及$f_{\beta C}$。计算重叠片段的$\langle f_{\alpha N}\rangle$，$\langle f_{\beta N}\rangle$，$\langle f_{n\alpha N}\rangle$，$\langle f_{n\beta N}\rangle$，$\langle f_{\alpha C}\rangle$，$\langle f_{\beta C}\rangle$，$n\langle f_{n\alpha C}\rangle$及$\langle f_{n\beta C}\rangle$。若$j_{\beta N}-j_{\beta C}$与$j_{\alpha N}-j_{\alpha C}$的交叉重叠发生在$\alpha$-螺旋区的C端或$\beta$-折叠区的N端且$(\langle f_{\alpha C}\rangle+\langle f_{n\alpha C}\rangle)>(\langle f_{\beta N}\rangle+\langle f_{n\beta N}\rangle)$则重叠区域为$\alpha$-螺旋结构，反之为$\beta$-折叠结构。若交叉重叠发生在$\alpha$-螺旋区的N端或$\beta$-折叠区的C端且$(\langle f_{\alpha N}\rangle+\langle f_{n\alpha N}\rangle)>(\langle f_{\beta C}\rangle+\langle f_{n\beta C}\rangle)$则重叠区域为$\alpha$-螺旋结构，反之为$\beta$-折叠结构。

5．输出

程序输出蛋白质的各二级结构单元在肽链中的位置(打印机列表输出)和/或输出其二级结构图示及肽链的疏水极性曲线图(绘图仪输出)。图9-4是其中的一例。

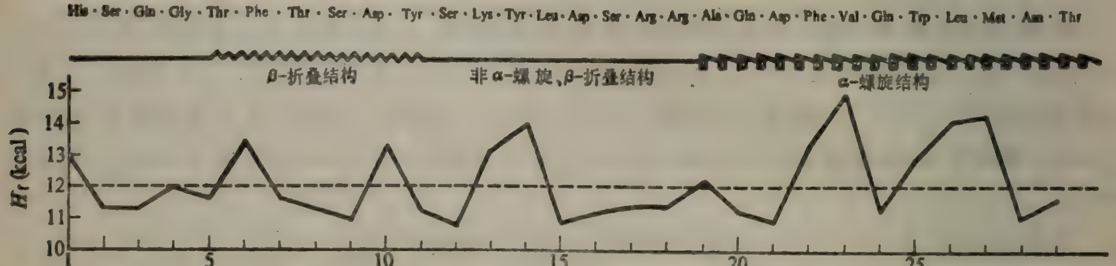


图 9-4 算法的输出形式

第一排为蛋白质的氨基酸顺序，第二排是其二级结构图示，曲线图是蛋白质的疏水极性图。横坐标是氨基酸位置，纵坐标为对应的疏水极性值(kcal)。

根据上述理论和方法，刘志平和江寿平编了用于TRS-80(I)微型机的FORTRAN IV 程序。并对细胞色素b_5硫氧还蛋白、腺苷酸激酶和磷酸甘油酸激酶的二级结构进行了分析。结果与X

射线、Chou和Fasman结果比较，由本算法预测出来的结果与X射线分析结果基本吻合。由于算法中用于决定函数$nf,nb,\langle P_\alpha\rangle$及$\langle P_\beta\rangle$之值的$P_s(R_i),f_s(R_i),b_s(R_i)$，同Chou和 Fasman 且基本遵守Chou和Fasman规则。故结果与Chou和Fasman结果也基本一致。然而，由于本算法可用计算机实现，故可对大分子蛋白质的二级结构进行快速预测。

由于本算法提供了以核苷酸作为输入的输入方式，故可在不知道蛋白质的氨基酸顺序的时候，从其核苷酸顺序推测出它所编码的蛋白质在二级结构方面的信息。

3．蛋白质二级结构预测例子[10-16]

(1) 乙型肝炎病毒的表面抗原蛋白分子

将乙型肝炎病毒的表面抗原(简称HBsAg)，注射到哺乳动物体内。可以引起抗体的合成。这些抗体可以和HBsAg专一地结合，实验表明，蛋白质抗原的抗原专一性不仅取决于分子表面的局部构象及抗原决定簇的性质和数目，而且依赖于它们的空间分布，即整个分子的空间构象。因此，研究 HBsAg 的空间构象，分析、比较不同亚型的 HBsAg 在构象上的异同，对于了解 HBsAg 免疫的分子机制，寻找理想的抗原专一性较强的蛋白片段或区域为设计合成疫苗提供重要依据。

随着 DNA 序列测定技术的发展，使可以不必分离、提纯一些含量极微且难以提纯的蛋白质而按密码字典间接地从核苷酸顺序推知其一级结构。根据蛋白质分子的高级结构是由其一级结构决定的这一实验事实和近年来发展，积累起来的有关蛋白质高级结构信息及基于这些信息而建立起来的各种蛋白质高级结构预测方法，就有可能根据蛋白质的一级结构很快地了解这些蛋白质的高级结构的一些信息。刘志平和江寿平[11,12]利用自己所编制的蛋白质二级结构预测程序，对 4 种亚型的 HBsAg-adr型[13]，adw2 型、ayw 型、adyw 型的二级结构进行了预测、比较。给出了它们的疏水、亲水极性强弱的曲线分布图，并对该抗原蛋白的抗原决定簇在肽链中的定位进行了预测。

抗原蛋白的抗原决定簇的定位：抗原决定簇的大小一般为 6—7 个氨基酸残基。一条抗原蛋白肽链可以有若干个抗原决定簇。对于球蛋白分子，抗原决定簇很可能位于肽链转折处，球蛋白分子表面突出的地方。这种区域多含有带电荷的，亲水性较强的氨基酸残基。因此，至少有某些抗原决定簇与肽链的亲水性大小有关。Hopp 等人认为肽链中对应于具有最大亲水性值的顺序为抗原决定簇所在或紧邻着抗原决定簇的区域。

程序输出各蛋白质肽链的亲水，疏水性分布图(图9-5，6)。图 9-6 中，横坐标表示肽链的氨基酸残基顺序位置，纵坐标为第 j 个氨基酸残基到第 $j+5$ 个残基这 6 个氨基酸残基顺序的平均极性值。图 9-6 中列出的氨基酸顺序为亲水性最强处，即抗原决定簇所在或紧邻着抗原决定簇。曲线上方的数值为相应处的亲水值。图 9-5 中，横坐标的意义同图 9-6，纵坐标表示相应的残基的疏水性值(H_i)。为清楚起见，我们只绘出了 adr 型的亲水性分布图，其它亚型与 adr 型的异同见表 9-6 。在图 9-5 中我们也只给出了 4 种亚型的部分顺序的疏水性分布图(150—180)。我们将在下面对这段顺序进行讨论。图 9-5，图 9-6 是通过一绘图子程序在与 TRS-80 微型计算机相连的绘图仪上输出的。

由表 9-5 可以看出，4 种不同亚型 HBsAg 的三级结构基本相似，β-折叠占相当高的比例，α-螺旋结构很少或几乎没有，由于个别地方氨基酸的差别，出现了二级结构的变化，例 156—172

表 9-5　4种亚型HBsAg的二级结构

二级结构	adr	adw2	ayw	adyw
β-折叠	12—28	12—28	12—28	12—28
	35—42	35—42	35—42	35—42
	47—52	47—52	45—49	45—49
	72—104	72—104	72—104	72—104
	112—116	114—118		
	121—126	121—126	121—126	121—126
	137—141	137—141	137—141	137—141
	146—150	146—150	146—150	146—150
	156—163	156—163		
	168—173	168—173		
	179—184	179—184	179—184	179—184
	189—201	189—201	189—201	189—201
	205—209	205—209	205—209	205—209
	218—226	218—226	218—226	218—226
β-转角	31—34	31—34	31—34	31—34
	44—47	44—47		
	68—71	68—71	68—71	68—71
	119—122	119—122	119—122	119—122
	128—131	128—131	128—131	128—131
	143—146	143—146	143—146	143—146
	151—154	151—154	151—154	151—154
	201—204	201—204	201—204	201—204
α-螺旋			160—170	160—170

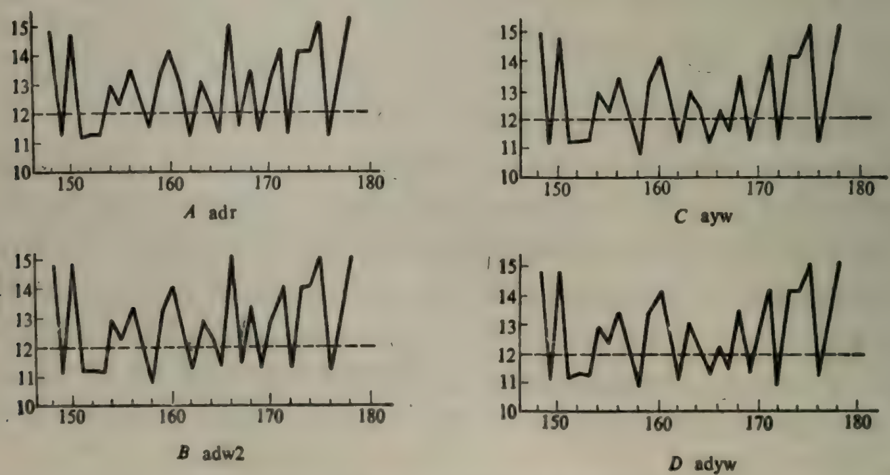


图 9-5　HBsAg 中片段 150—180 的疏水极性分布图
图中 A、B、C、D 分别为 4 个亚型的 HBsAg

表 9-6

亚　型	亲水极性最强处	亲水极性值
adr	141	0.783
adw2	141	0.967
ayw	141	1.08
adyw	141	1.08



图 9-6　adr亚型的HBsAg的亲水极性强弱分布图

处，从它们各自的疏水性曲线(图 9-5)可以看出，160—166处，其疏水性强弱显周期为 2—3 个残基变化，166—170 处，其疏水性显周期为 1 个残基变化。adr、adw 2 二亚型第 168 个残基为Val，而ayw、adyw 二亚型相应处为 Ala，由于 Val 比 Ala 具有大得多的疏水性且 Val 具有较高的β-折叠势，Ala 具有较高的α-螺旋势，导致区域 168—170 的$\langle P_\alpha\rangle$与$\langle P_\beta\rangle$差别很大。故对于 adr、adw 2 亚型，168—170 为β-折叠结构，而 ayw、adyw 亚型为 α-螺旋 结构。尽管 adr、adw 2 二种亚型的 164—167 顺序被预测为无规结构，但从其疏水曲线看。可能这段区域含 有较弱的 α-螺旋结构。

由图 9-6 及表 9-6 可以看出，4 种亚型蛋白分子的亲水性最强者 都在 141—146 处，进一步考察它附近区域的二级结构可以发现，137—141，146—150 为β-折叠结构，143—146 为β-转角结构。因而片段 137—141 很可能 借 143—146 这个 β-转角与顺序 146—150 形成 反平行 β-折叠片，并通过二个区域中 的半胱氨酸(Cys)或其它区域中的 Cys 形成的二硫键 及其它相互作用而得以稳定。这与实验观察得到的关于球蛋白抗原决定 簇的构造的一般结论[14]是一致的。

蛋白质抗原的高级结构阐明及其氨基酸残基的极性分析，对于寻找 理想的抗原专一性强的抗原性区域具有明显意义。Hopp[15]利用 亲水性 分析在 HBsAg 中找到了一段具有 最强亲水性的顺序并合成了一段类似物。实验表明这段顺序是有一定的免疫活性的。

要彻底阐明蛋白质抗原的抗原专一性的分子机制，仅了解其二级结构是不够的。蛋白质的抗原专一性更多情况下依赖于它的整个分子的空间构象，三级结构的再折叠，卷曲可以使某些抗原决定簇位于球蛋白分子的内部，而使另一些抗原决定簇位于分子的表面。从 4 种亚型的 疏水性分析可以看出，79—99，190—201，218—226 这三段区域里的氨基酸几乎全是疏水性 氨基酸残基，它们可能可以和肽链中其它一些疏水残基相互 作用形成一疏水核。β-折叠区 168—173 其疏水性强弱显周期为一个残基的变化，可能位于球蛋白表面。

(2) 干扰素

干扰素(IFN)对细胞功能可产生各种影响，尤其是它的抗增殖作用及抗病毒活性，引起了人们的很大兴趣。IFN 分为 α、β 和 γ 三类，由圆二色性(CD) 谱估计它会有 55—70% 的螺旋。Hayes[16]应用 Chou-Fasman 方法估计了 Hu-IFN-α 约有 55%的 α-螺旋，16%的β-折叠，19%

的无规卷曲和1%的转角结构。Wetzel[17]等用改进的Garnier等的方法计算IFN-αA，它含有65%的α-螺旋，很少β-折叠，结果表示于图9-7中。程玉华[18]等也对IFN的二级结构进行了一些分析。王良鹏和江寿平应用自编的蛋白质二级结构预测程序预测了IFN-γ的二级结构和它的疏水性和亲水性分布图(表9-7，图9-8)。

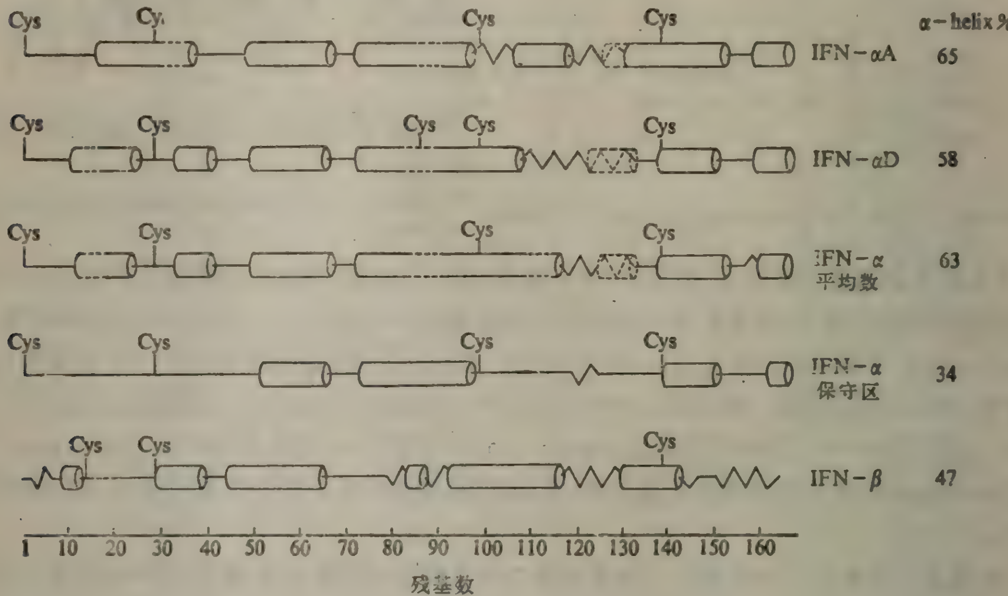


图 9-7 Hu-IFN-α 和 β 预测的二级结构

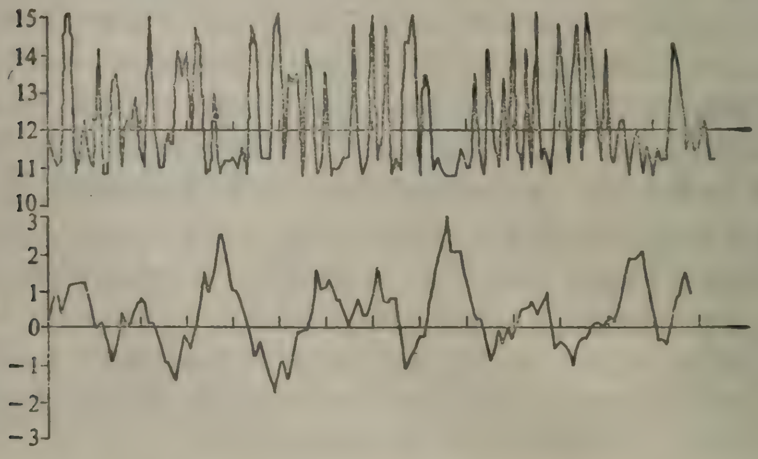


图 9-8 IFN-γ 的疏水性(上)和亲水性图(下)

9.1.2 核酸分子二级结构的预测

随着DNA顺序测定方法的不断改进提高以及基因工程的迅速发展，迄今为止已经积累了

表 9-7 IFN-γ的二级结构

α-螺旋	4—13，34—58，64—82，86—89，102—121
β-折叠	27—33，97—101，133—137

大量的 DNA 序列资料[19,20]。自 1977 年测定了第一个完整的 DNA 分子噬菌体 $\phi\times174$ DNA 序列后，人们已经陆续测定了细菌质粒 pBR322，动物病毒 SV40，噬菌体 T7，fd，f1，MB 等一系列的 DNA 序列。但是，对已有的包含极其丰富信息的核酸序列资料的分析、处理和研究，单靠人力是难以胜利的，而必须借助有针对性的方法和计算机进行分析处理，为此人们编写了各种计算机程序，它们包括下列一些功能[20-30]：(1) 它能方便地实现对核酸序列资料的输入、存贮，删改，编辑，格式化处理以用于核酸序列数据文件的建立和管理。(2) 能通过对核酸序列数据的复制以实现 DNA，RNA 的互补链的构造，存贮与序列数据文件资料的打印输出。(3) 它能通过对核酸序列资料的检索及多种处理以实现 DNA、RNA 链的氨基酸密码翻译，氨基酸密码子的频度计算及检索，DNA 分子的核酸限制性内切酶图谱的构造，任意指定子序列的检索和发夹结构的识别与预测。(4) 它能通过对核酸序列进行比较以实现序列中重复序列的检测及对存在重叠序列片段的二核酸序列实现拼接组合等多种功能。

必须指出，随着分子生物学，基因工程、核酸序列分析技术广泛深入的发展，测定的核酸序列愈来愈多，这就要求我们通过计算机对大量序列数据进行分析研究了解核酸分子结构信息和动态过程，通过计算机来模拟和预测它们的高级结构，为核酸分子的设计提供有益的信息。这方面工作近些年来引起人们的高度重视，进行了许多研究[20,21]。

9.1.2.1 RNA 大分子的二级结构预测方法

我们知道，单链 RNA 分子通过自身的回折可以形成部分螺旋，突环相间排列的高级结构。最近十年来 Tinoco[31]，Nassinov[32]和Zuker[33]等人进行了一系列研究，建立了多种预测方法，并在实际问题中得到应用。乐树云和江寿平[34,35]提出改进的预测算法。算法是在具有最大碱基配对预测算法的基础上发展起来的，它的基础是拓扑平面的最大C-匹配原理和现有的单链RNA折叠构象的热力学数据。它能方便地扩充而耦联与单链核酸分子折叠构象有关的其它信息。在算法中我们结合了 Nussinov 算法的数学描述的严谨性和 Zuker 算法中引入的“面权重”处理的有效和灵活的长处。同时为了减少计算量我们在计算图中各子图的权重时提出了一系列新的算法公式，并在数据结构方面作了一定技术上改进。

1. 拓扑平面图的最大 C-匹配处理：对于一长度为 n 个核苷酸残基的单链核酸分子，通过碱基的配对可以折叠成形形色色的螺旋，突环相间排列的二级结构。但从生物意义的角度而言，令人最感兴趣的是该分子的最可能的折叠结构，即它的具有最小生成自由能的折叠结构。核酸分子的折叠结构的生成自由能的来源有二个方面，一方面是来源于结构中螺旋区域的氢键能和碱基间的堆积能，另一个则来源于成环区域内，发夹环，内环以及凸环结构的生成自由能。因此单链核酸分子的折叠结构的总生成自由能即为各部分结构的生成自由能总和。在算法中，引用了 Saler 汇编的单链核酸分子各折叠结构的生成自由能的热力学参数来计算各种碱基的堆积能

(包括了氢键能)及各类环结构的生成自由能。对于单链 DNA 的折叠结构的生成自由能，是参照单链 RNA 的折叠结构 的热力学参数。

为了寻求单链核酸分子的最小自由能的折叠结构，采用了拓扑平面图的数学处理。即在处理中用一重图$G(X,E)$表示一个由 n 个核苷酸残基的单链 RNA 分子 的折叠 结构。$G(X,E)$中X $\{x_1,x_2\cdots,x_{n-1},x_n\}$，即图中各顶点的集合，$E$ 为图中连接二个顶点的弧的集合。图$G(X,E)$和折叠结构之 间具有以下的一一对应关系。 图中一顶点 X_i 对应于 RNA 分子中第 i 个核苷 酸残基(序号是按 5′端到 3′端的方向依次编号)，核苷酸链的磷酸二酯键则用图中的弧表示，而图中任意不相邻的棱则用来连接链中能相互配对的碱基。因此经拓扑平面图的数学处理后，寻求具有最小生成自由能的单链 RNA 分子的最稳定的结构就可通过寻求该对应的拓扑平面图的最大C-匹配来解决。应该说明的是，预测具有最大碱基配对数折叠结构的算法 仅仅是求该拓扑平面图的最大 C-匹配方法的特款，即在最大匹配处理中拓扑平面中两非相邻顶点 x_i 与 x_j 相 匹配 时取权重 $E_{ij}=1$ 而在最大 C-匹配处理中 E_{ij} 的值取决于 拓扑平面图中与棱 $x_i x_j$ 相关的 面的结构类型。因此现在的问题就是寻求该拓扑平面图中具有最大权重(即具有最小生成自由能)的所有不相邻的棱的集合。

2. 扑拓平面图 $G(X,E)$中面与棱

现有一 RNA 片段序列 5′C-C-A-C-A-U-G-C-A-A-G-C-A-A-U-G-A-A-A-G-G-A-G-G-U-U-A-C-U-A-A-A-C-A-U-A-U-U-A-G-C-C-A-C-C-A-C-C-U-U-C-U-U- A-U-G-G 它可以折叠成如图 9 -9所示的结构，与它对应的拓扑平面图 $G(X,E)$如图 9 -10 所示。按照上述的处理有：RNA 分子中57 个核苷酸残基依次对应于图 $G(X,E)$中 57 个顶点$x_1,x_2\cdots,x_{56},x_{57}$，RNA 分子中磷酸二酯键对应于图 $G(X,E)$中的弧，而 RNA 分子所形成的 氢键则用图 $G(X,E)$中互不相邻的棱表示，这里我们作了如下的约定：(1) 每个顶点至多有一条棱使它与其相匹配的顶点相连，而棱与棱是互不相交的。(2) 拓扑平面图中的面定义为由弧及 棱所 围成的一个区域 区域内任意二点能由不与弧(棱)或顶点相交的连续曲线相连接，即拓扑平面图的面内既不含图的顶点，也不含图的弧(及棱)。各类面定义如下：

(1) 边界上只有一条棱的面称为发夹环(图 9 - 9，9 -10 中用字母 H 表示的区域)。

(2) 边界上有二条棱和二条弧的面称为螺旋堆积区(图 9 - 9，9 -10 中 用 字 母 S 表示的区域)。

(3) 边界上有二条棱和三条弧的面称为凸环(图 9 - 9，9 -10 中用字母 BU 表示的 区域)。

(4) 边界上有二条棱和三条弧以上的弧的面称为内环(图 9 - 9，9 -10 中用 字 母 I 表示 的区域)。

(5) 边界上有三条和三条以上棱的面称为分叉环(图 9 - 9，9 -10 用字母 BF 表示的区域)。

有了上面的约定，根据拓扑平面图的最大 C-匹配原理，就可以建立一个递归的，动 态 的，快速的算法，并以 FORTRAN 语言编成计算机程序。应用这个程序对免疫球蛋白 $\gamma 1$ 重链的 mRNA 片段序列，大肠杆菌 16 SrRNA 片段，以及脊髓灰质炎病毒 RNA 片段序列的二级结构进行了预测，其结果如下：

(1) 免疫球蛋白 $\gamma 1$ 重链的 mRNA 片段序列的二级结构

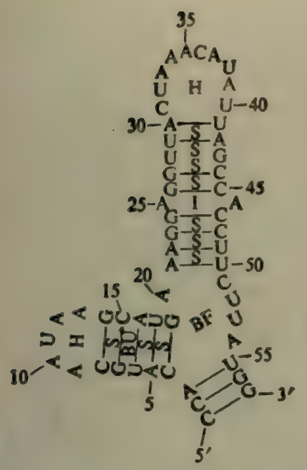


图 9-9　57 核苷酸的 RNA 片段序列的可能的折叠结构

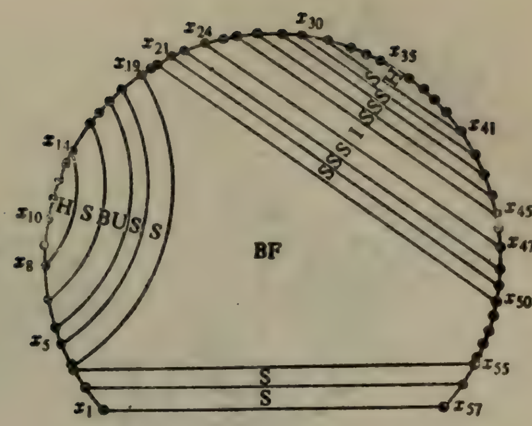


图 9-10　与图 9-9 折叠结构相应的拓扑平面图 $G(Z,E)$

为了同已有的算法所作的 RNA 分子的二级结构预测的结果进行比较，我们对免疫球蛋白 γ1 重链的 mRNA 片段序列(459个核苷酸残基)进行了计算机预测。根据上述的算法在 IBM4341 计算机计算结果表明：该片段序列能折叠成具有 −194 kcal/mol 最小生成自由能的折叠结构，如图 9-11 所示。在该预测的结构模型中共有 126 个碱基对，片段序列中能配对的碱基的比率约为 55%（已包括碱基 G 和碱基 U 的配对）。该预测的结构模型同 Zuker 和 Stiegler[33]所预测的结构模型相比较表明约有 95 对碱基对是相同的。占整个螺旋区域的 76% 左右。在图 9-11 配对区域中直线的螺旋区域即为二种算法预测的一致的螺旋区域。显然这里预测的免疫球蛋白 γ1 重链的 mRNA（459 核苷酸）片段序列的二级结构的基本特征是类似于 Zuker 和 Steigler 计算的结果。

（2）大肠杆菌 16 *S* rRNA 片段序列(567—833)的二级结构模型。

Noller 和 Woese[36]在 1981 年基于多种方法研究，提出了大肠杆菌 16 *S*rRNA 分子的二级结构模型(图 9-12)，以后他们又作了些改进。根据这个模型，大肠杆菌 16 *S*rRNA 的二级结构大致可以分成三个区域：5′末端域，中心域和 3′末端域。为了便于同现有的 RNA 大分子二级结构模型相比较，选取了大肠杆菌 16 *S*rRNA中心域的片段序列(567—883)，用上述算法进行了计算机预测，所预测的二级结构模型如图 9-12 所示。该片段序列能折叠成具有 −142.16 kcal/mol 最小生成自由能的稳定结构。该结构模型中共有 95 对碱基对，配对碱基比率约为 60%。这一比率基本上与物理方法预测的结果(60%—70%)相同。根据 Noller 和Woese提出的结构模型，该片段序列能折叠成共 85 对碱基对的螺旋区域，配对碱基比率约为 54%。经比较表明，该折叠结构中约有 86%的螺旋区域与乐树云和江寿平预测的结构模型是一致的。在图 9-12 中画直线的螺旋区域即为二种结构模型的共同的螺旋堆积区域。由此可见，就大肠杆菌 16 *S*rRNA 中心域片段序列(567—883)而言，由上述算法所预测的二级结构模型基本上同 Noller 和Woese基于多种方法研究结果所提出的模型是很类似的，它们结构的基本特征是一致的。Noller 和他的

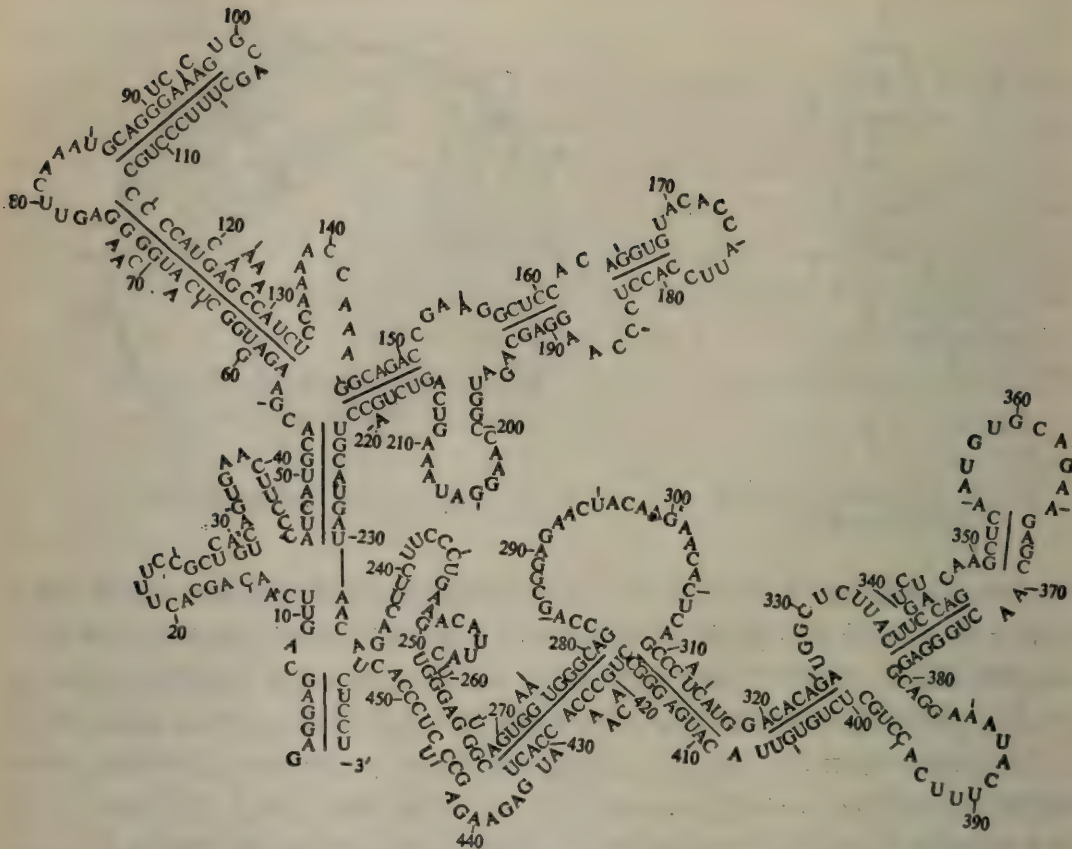


图 9-11 免疫球蛋白 γ1 重链 mRNA 片段序列(459 核苷酸残基)的具有最小生成自由能的折叠结构

同事又预测了 23S rRNA(约 90 个残基)的二级结构[37]。

(3) 脊髓灰白质炎病毒 RNA 片段序列(1—740)的二级结构模型

脊髓灰白质炎病毒 RNA 约 7433 核苷酸长,它的结构基因具有典型的高等动植物病毒 m-RNA 的特点。它的 5′端有一较长的不翻译区(1—740),这个区域与基因表达调节功能有关。根据上述算法所预测的该片段序列(1—740)的二级结构如图 9-13 所示。所预测的结构模型表明该 RNA 片段序列能折叠成－304.89 kcal/mol 最小自由能的树状稳定结构。在该稳定结构中约有 228 对碱基对,彼此能配对的碱基所占的比率为 62%。几个单链环区中彼此之间存在着潜在配对的可能性有可能进一步折叠成更高层次的结构(如图9-14 中 293—30)发夹环中 AACC-(3′)与 542—548 发夹环中 UUGG-(3′),142—149 发夹环中 AAA-(3′)与 559—569 单链环中 UUU-(3′)以及 118—127 单链环中 AGUUC-(3′)与 616—620 单链环中(3′)-UUAAG 相互配对而形成左右折叠,而 420—428 发夹环中AUUGA-(3′)与 708—728 发夹环中(3′)-UAACU以及 453—479 发夹环中 UAAU-(3′)与 646—654 发夹环中 AUUA-(3′)相互配对而上下折叠,进而形成更紧密的更高层次的空间结构。

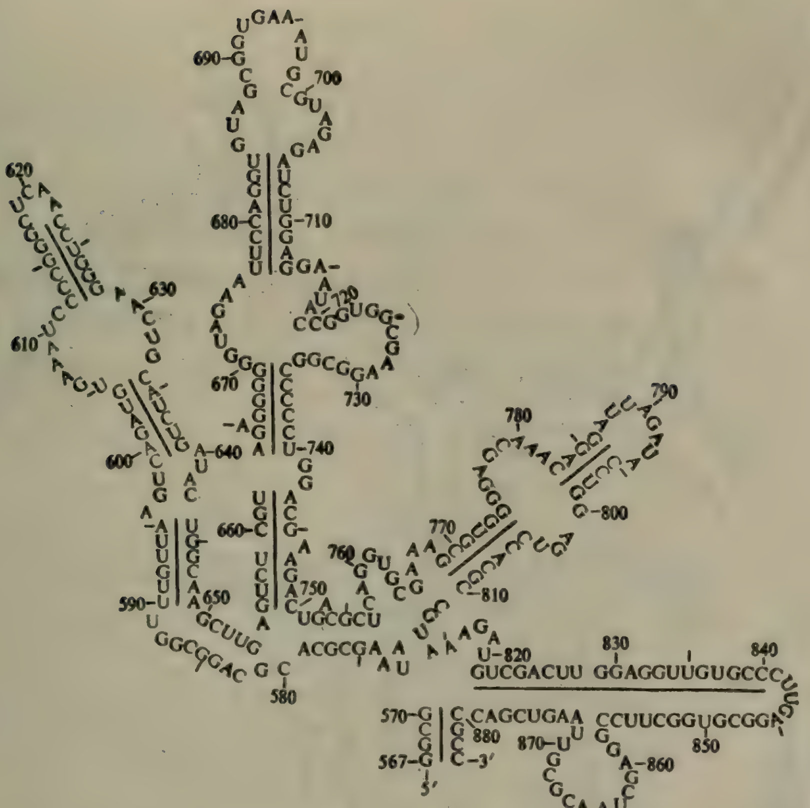


图 9-12 大肠杆菌 16 *S* rRNA 中心域(567—883)片段序列的具有最小生成自由能的二级结构

总之上述算法对于预测 RNA 分子具有最小自由能的稳定二级结构是有效的。利用该算法已研究了5.8*S* rRNA 分子的二级结构模型[38],mRNA 翻译起始区域结构特征与翻译效率间的关系[39],和大肠杆菌 RNA 聚合酶识别的 54 个启动基因的结构特征研究[40]以及 RNA 分子折叠的动态过程[41]等。得到了十分有益的信息。可以相信随着算法耦合实验方法所得的 RNA 分子结构的特征信息的不断完善,利用计算机预测 RNA 大分子的二级结构是切实可行的。

9.1.2.2 5.8 *S* rRNA 的二级结构[38]

真核生物 5.8 *S*rRNA 是细胞核糖体中的大亚基的结构成分,到目前为止约有二十几种 5.8 *S*rRNA 分子的序列被测定。现有的实验已表明 5.8 *S*rRNA 分子对于真核生物核糖体的活力表现和它的结构起着重要作用。为了研究 5.8 *S*rRNA 的结构和功能,对于它的二级结构已基于多种方法提出了一些可能的模型(Nazar[42]等人,Ford和Mathieson[43],Luoma和Marshall[44]),但迄今为止尚没有一致公认的二级结构模型。刘次全等计算 5 *S*RNA[45],乐树云和江寿平对脊椎动物、高等植物、非脊椎动物和真核类微生物共 13 种 5.8 *S*rRNA 分子,借助于单链 RNA 分子二级结构预测算法,通过计算机的处理提出了它们的具有最小生成自由能的二级结构模型。并对这些模型及它们的高度保守序列在其结构模型中呈现的结构特征进行了分析和讨论。

脊椎动物的 5.8 *S* rRNA 二级结构:来源于脊椎动物的 7 种 5.8 *S* rRNA 分子的二级结构

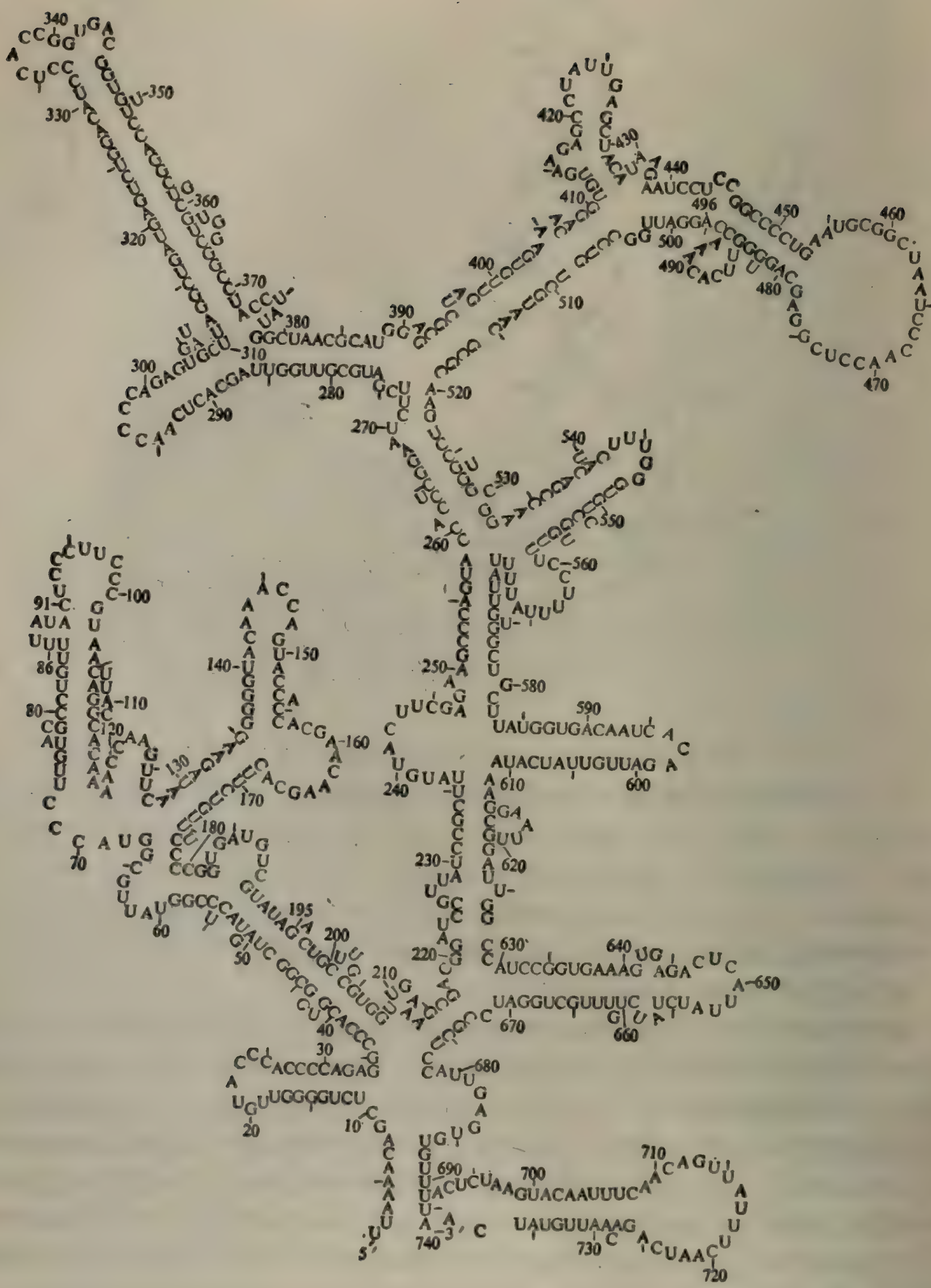
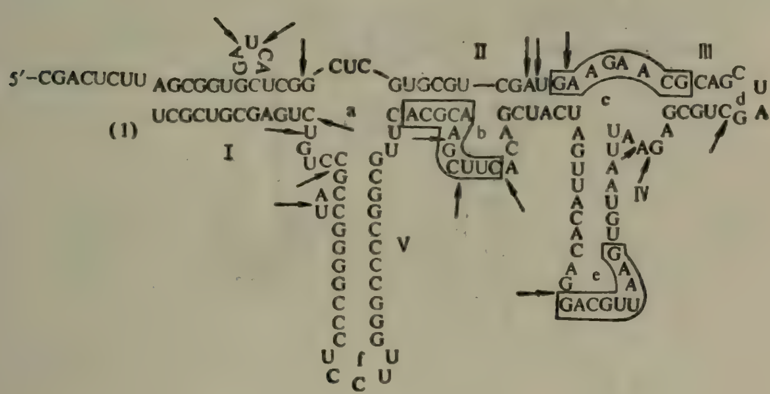


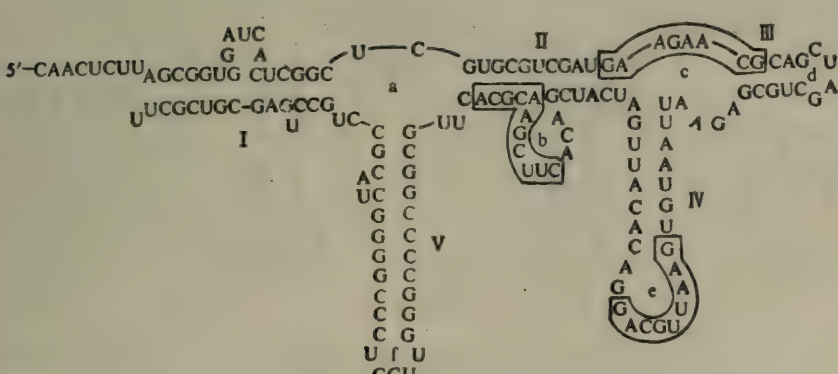
图 9-13　脊髓灰白质炎病病毒 RNA 片段序列(1—740)具有最小生成自由能的二级结构

模型分别显示在图 9 -14 *A*—*E*。这 7 种不同的 5.8 *S*rRNA 分子各自折叠的最稳定的二级结构模型是很相似的。这些模型同 Nazar 等人提出的 5.8 *S*rRNA 分子的二级结构模型是相近的，而与 Luoma 等人、Ford 等人提出的二级结构模型有较大的差异。这里我们引用了 Nazar 等人的核糖核酸酶水解实验依据，如图 9-14 *A* 和图 9-14 *C* 所示的箭号位置。可见具有最小自由能的折叠结构模型与 Nazar 等人的实验结果基本相符。在这 7 个 5.8 *S*rRNA分子的二级结构中，它们的共同的特征是结构中有相对稳定的 5 个螺旋臂和 6 个单链环区。其中螺旋臂 *IV* 是富 A-U 配对的螺旋区域，而螺旋臂 *V* 则是富 G-C 配对的螺旋区域，且在它们的上方都对称地折叠着一个单链环。同时值得指出的是它的 3 个环区 b，d 和 e 环恰巧是它们 3 个强保守序列所处区域。

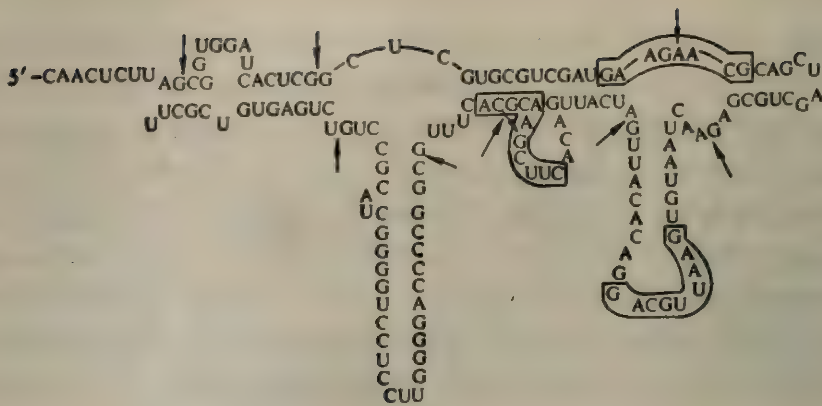
可见由于它们的亲缘关系相近，而使二级结构呈现共同的特征。根据这些模型可以提出一个对来源于脊椎动物的 5.8 *S*rRNA 分子更为一般的结构模型（图 9 -15）。它的主要特征参数参见表 9 - 8。



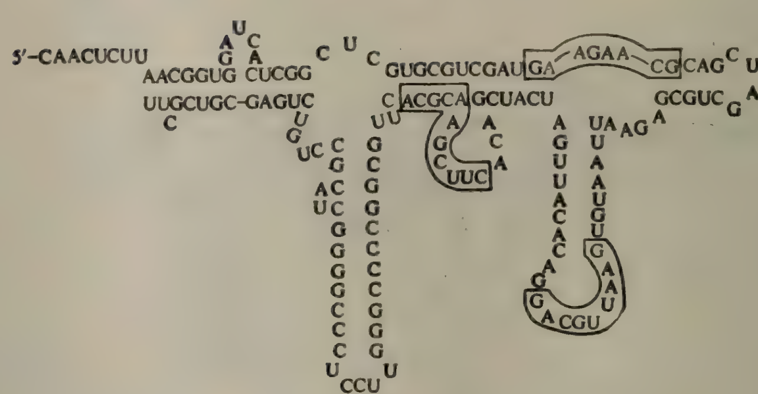
A



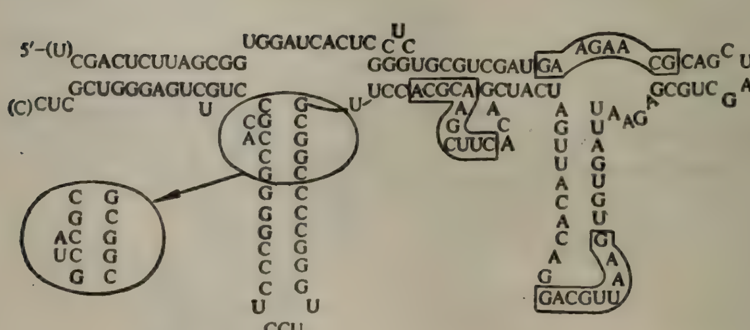
B



C



D



E

图 9-14 *A*. Novikoff 腹水瘤细胞 5.8 *S*rRNA 二级结构。其中方框内的残基为它们的保守序列，图中箭号指示了核糖核酸酶作用敏感的位置，鸡胚细胞 5.8 *S*rRNA 二级结构模型除序列中第 2 个（以5′-端起计数）残基为 A 取代图中的第二位残基 G 外与该模型图完全一致，且在 3′端增加一残基 U

B. 小鼠(MPC-11 细胞)5.8 *S*rRNA 二级结构

C. 虹鳟鱼(RTG-2)5.8 *S*rRNA 二级结构箭头号指示了核糖核酸酶(RNaseT)作用敏感的位置

D. 甲鱼心脏细胞(CCL 50)5.8 *S*rRNA 二级结构

E. 爪蟾和北方蟾体细胞 5.8 *S*rRNA 二级结构

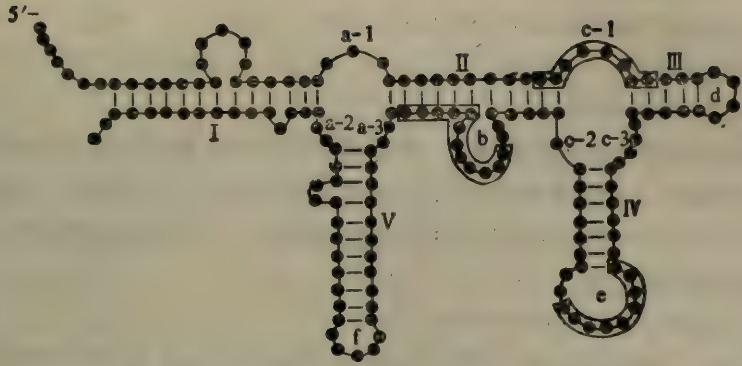


图 9-15 脊椎动物 5.8 SrRNA 二级结构的一般模型

表 9-8 脊椎动物5.8SrRNA二级结构一般模型中结构参数

螺旋臂	双螺旋片段数	碱基对数	单链环区	碱基数
I	1—3	12—13	a-1	2—10
			a-2	0—4
II	2—3	12—14	a-3	1—2
			b	9
III	1	6	c-1	4
			c-2	0—1
IV	1	7—8	c-3	4
V	2	11—13	d	4
			e	10
			f	3—5

9.1.2.3. 大肠杆菌 RNA 聚合酶识别的启动子的结构特征

原核基因在细胞中的表达能在 RNA 转录水平上被控制，这是由于 RNA 合成过程中位于 DNA 序列中若干调节信号作用的结果。位于 DNA 转录起始区域的启动子就是通过与 RNA 聚合酶的相互作用来实现转录水平上的调节。已有的实验结果表明位于这些特殊区域的片段 DNA 序列的特征在很大程度上决定了 DNA 被转录的效率。现有的研究在一定程度上也表明启动子在它们的一级结构上都应具有一定的共性，即在 mRNA 合成的起始位点前－35 处的 TTGACA 和－10 处的 Pribnow Box 的序列保守性，以满足启动子与大肠杆菌 RNA 聚合酶相互作用的需要。同时它们之间在结构上也必然存在着一定的差异，而正是这些差异可能反映了不同的启动子各自在功能上的差异。但是它们的高级结构究竟具有什么特征呢？这些结构特征与功能上的差异究竟如何发生并联，还是研究得很不充分的。乐树云和江寿平根据能被大肠杆菌 RNA 聚合酶识别而相互作用的 54 个启动子[46]为研究对象，分析了它们可能潜在的高级结构模型，它们的－35，－10 区域的特征序列在这些高级结构模型中的结构特征，并试图从这些结构特

征与它们功能上的差异发生关联。

研究的对象是可被大肠杆菌 RNA 聚合酶识别的 54 个启动子。它们有二个保守区域，即位于 mRNA 合成起始位点前－10 区域的 TATAAT 和位于－35 区域的 TTGACA。

1．启动子的折叠结构模型

借助于计算机对启动子序列的处理[35]，可以得到它们的最稳定折叠结构模型。若以双链考虑，它们的结构可呈镜象对称的折叠构型，它们的部分结构如图 9-16 所示。由图 9-16 可知，这些启动子的最稳定折叠结构模型都呈现了基环结构的特征。可以认为启动子所呈现的基环结构的构型是它的高级结构的共性。

2．保守序列在基环结构中的结构特征

根据上述预测的折叠构型和启动子的共同特性，分析了启动子中二个保守区域在其折叠结构中的结构特征，可以看到在这 54 个启动子中除了 galPI 和 tet 外，它们的－10 区域即与大肠杆菌 RNA 聚合酶紧密结合区域恰巧都位于它们各自的折叠结构中的单链区域，或是单链与螺旋基的结合区域。同样除了 λcin，$\lambda R_{R'}$，bioB，rpoB 和 rrnEI 等启动子外它们的－35 区域即大肠杆菌 RNA 聚合酶能识别的信号域也均处于单链环区域或环基的接合区域。启动子在高级结构上显示的这种结构特征是有利于大肠杆菌 RNA 聚合酶对转录起始区域的识别和它与启动子的相互作用的。我们知道原核生物 RNA 转录的起始过程机制是 RNA 聚合酶首先与启动子中－35 区域的识别信号相互作用，再与启动子中－10 区域的 Pribnow Box 中的结合位点相结合后移到转录起始位点引发 RNA 的合成。因此启动子必须具有一定的空间构型适合它与 RNA 聚合酶的识别和相互作用。我们提出的启动子潜在的最稳定折叠结构所表现的基环结构的共同结构构型特征，以及启动子的－35、－10 区域中特征序列在基环结构中所呈现的共同结构特征，这在空间结构上有利于原核细胞 RNA 的这一转录机制的实现。同时我们也注意到已知的具有高启动效率的启动子如 λP_L，LacUV5，trpE.C.等它们的－10 和－35 区域中保守残基几乎都位于其基环结构中的单链环上(参见图 9-16)。由热力学理论可知这种结构特征所产生的空间约束是极有利于它们与 RNA 聚合酶的识别结合的。对一些启动子起始反应的动力学分析已指出，转录的起始速率与起始复合物生成的速率相比通常是相当快的。因此，启动子的效率取决于它的起始复合物生成速率。由于启动子中－35 和－10 区域在空间结构上的结构特征所产生的空间的约束对于转录起始复合物的生成是重要的，因此可以认为启动子中这二个保守区域在基环结构中的结构特征是与启动子的启动效率是密切相关的。

3．突变

现有的研究表明，启动子中，尤其是在它的－10，－35 二个保守区域中若发生碱基突变将会影响启动子的效率。有的突变增强启动子的启动效率(简称正效应)，有的却减弱启动子的启动效率(简称负效应)。为了探求其规律，我们根据启动子中碱基突变对启动子效应发生影响的实验事实，综合和分析了启动子中－10 和－35 区域中保守残基在其基环结构中的结构特征及在碱基突变发生前后的结构变化(表 9-9、图 9-16)，可以指出当发生正效应突变时(如 λP_{RM} 中第－32 位的残基 T 被 C 取代，LacI 中第－36 位的残基 T 被 C 取代)，特征序列中部分残基的突变引起了该区域的结构特征变化，即在启动子序列折叠的稳定的基环结构中，该保守区域特征序列所折

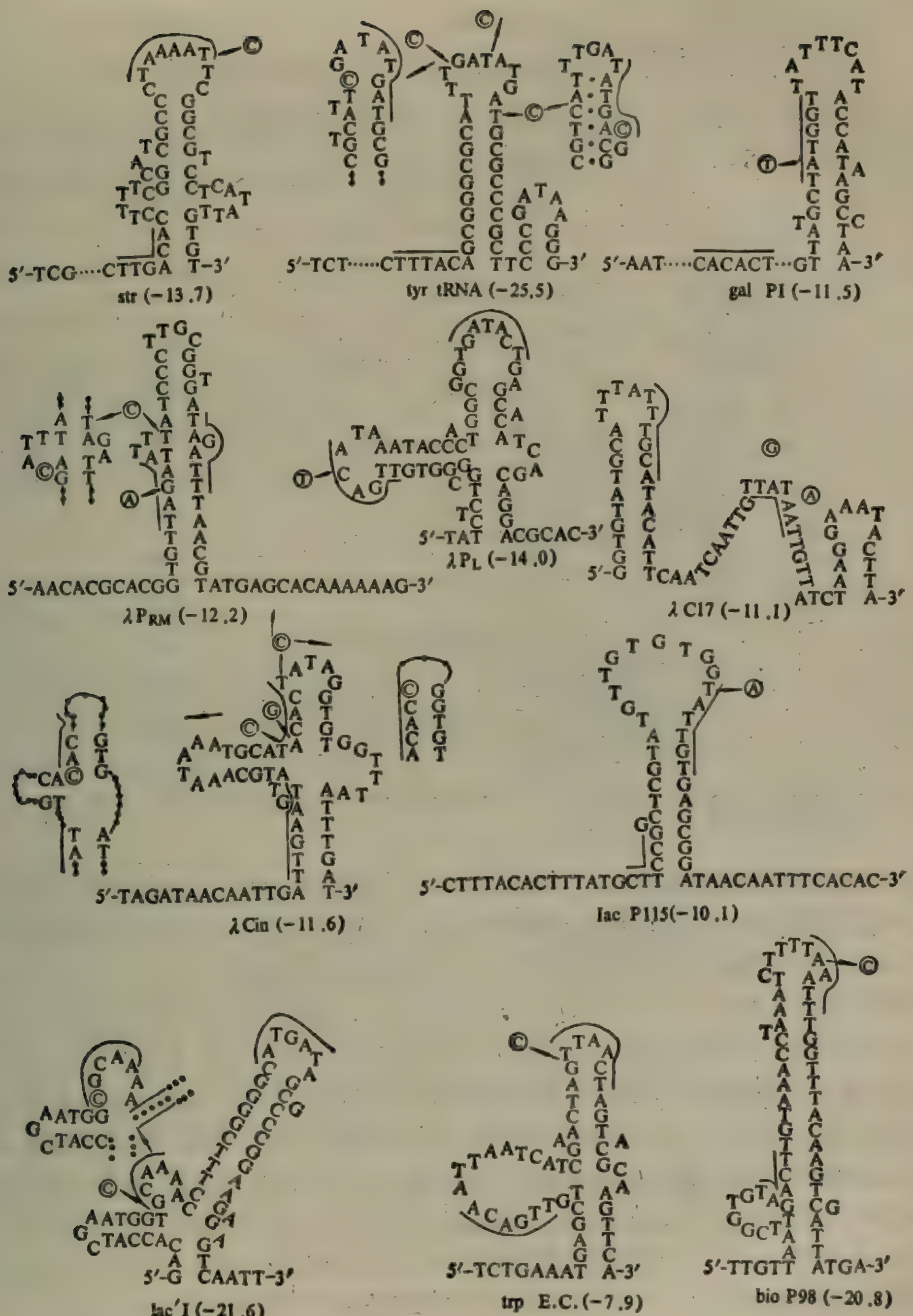


图 9-16　启动子中突变及其折叠结构

突变时碱基变化由圈号内的碱基所示，箭号指示突变的位置。括号内的数值为生成自由能值，单位 kcal/mol，划线区域为 -35、-10 区域中保守序列。双箭号所示的区域结构为突变后变化的构型。

表 9-9 启动子中碱基突变以及突变对－10，－35区域中结构特征的影响

启动子名	取代碱基	突变位置	－10区域结构特征	－35区域结构特征	启动效率
λP_{RM}	A⇒G	－34	茎结构的稳定性加强	茎结构的稳定性加强	下降
λP_{RM}	C⇒T	－32	内环中残基个数增加1（茎结构的稳定性减弱）	内环中残基个数增加1	上升
λP_L	T⇒C	－32	无变化	无变化	下降
λC17	G⇒A	－12	无变化	无变化	下降
λC17	A⇒T	－8	无变化	无变化	下降
λcin	C⇒T	－13	茎结构的稳定性加强	无变化	下降
λcin	G⇒A	－12	茎结构的稳定性减弱	无变化	下降
λcin	C⇒T	－8	茎结构的稳定性加强	无变化	下降
lac P115	A⇒T	－13	无变化	无变化	下降
lacI	C⇒T	－36	无变化	（环中残基数增1）茎结构的稳定性减弱	上升
trp E.C.	C⇒T	－13	无变化	无变化	下降
bioP98	C⇒A	－12	无变化	无变化	下降
str	C⇒T	－8	无变化	无变化	下降
tyr tRNA	C⇒T	－16	茎结构稳定性加强(环中残基数减少1)	无变化	下降
tyr tRNA	C⇒T	－13	无变化	无变化	下降
tyr tRNA	C⇒T	－8	茎结构稳定性加强(环中残基数减少2)	无变化	下降
gal P1	T⇒A	－12	无变化	无变化	下降

A⇒G，表示碱基A取代原碱基G。

叠的基环结构的稳定性减弱，而增加了单链环区中保守的残基。就 λP_{RM} 而言同时也引起了－10区域在启动子的基环结构中所表现的结构特征变化。而特征序列这一结构特征的变化是有利于RNA 聚合酶对启动子的识别和相互作用的。当发生负效应突变时其－10 和－35 区域中特征序列的结构特征多数没有发生变化(参见表 9-9、图 9-16)，部分的其保守区域中特征序列折叠的螺旋基环结构的稳定性增强而不利于启动子与 RNA 聚合酶的相互作用。其中仅有一例是例外，如启动子 λcin 在－12处的突变引起螺旋基环结构的稳定性减弱而该突变却起着负效应。虽然启动子中发生的突变对于它的效率影响的机制是复杂的，但是由于保守残基的突变引起保守区域在基环结构中结构特征变化，很可能就是诸种因素中的一个因素。我们认为，突变引起的启动子中特征序列的结构特征中基环结构的稳定性减弱是有利于RNA聚合酶和启动子间的相互作用的。

总之启动子的－10 与－35 区域中的特征序列在启动子的基环结构中所表现的结构特征是与启动子的效率有关的。这二个保守区域中保守残基是否处于单链环区或者是否处于不稳定的基环结构的结合区域，这可能和 RNA 聚合酶与启动子的识别和结合的效应是密切相关的。可以展望借助重组 DNA 技术，人工合成基因片段和计算机技术以改造启动子的特征序列在其高级结构上的结构特征而实现高效地转录也是可能的。同样可改造翻译起始的控制区域 DNA 序列，

而使该区域能折叠成最理想的构象状态而提高翻译的频率达到高效的表达。

9.2 蛋白质和核酸的构象和相互作用[47-49]

我们知道，一个多肽是以—NH—CHR—CO—形式的氨基酸残基连结在一起形成。而把 C^{α}—CO—NH—C^{α} 称之为肽单元，它形成一个平面基团，它的键长和键角表示于图 9-17 中：

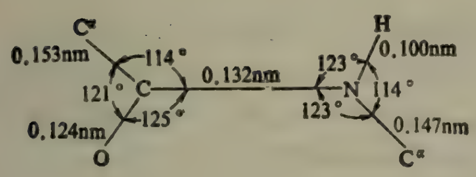


图 9-17 肽单元的标准大小

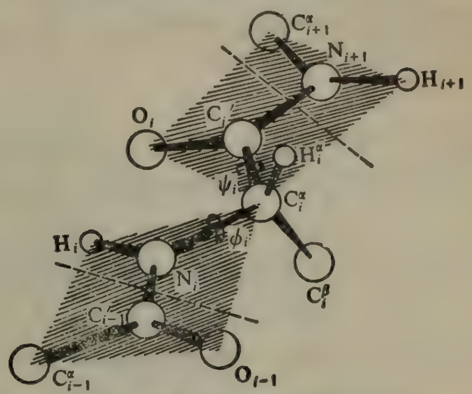


图 9-18 α-L-多肽链的透视图(反式，$\phi=\psi=180°$)

肽单元有顺式和反式两种，实验证明开环的多肽链中总是反式，而在环肽中可能出现顺式。两肽单元以 C^{α} 原子相连的两个键，N—C^{α} 和 C^{α}—C′ 是单键的特性，所以可以绕这两个键自由转动，这转动角各表为 ϕ 和 ψ，称之为二面角(图 9-18)，它们不同的值代表着不同的构象。由于不同原子之间允许接触距离的限制，所以 ϕ 和 ψ 仅能取一些值，因而仅一些构象是存在的。Ramachandran 把一个多肽允许构象的 ϕ 和 ψ 值表示在以 ϕ 和 ψ 为坐标的图上，称之构象图（图 9-19）。图中实线区域表示正常区域，虚线区域表示最大允许区域，(α_L)和(α_R)各为左手和右手 α 螺旋。(↑↑)和(↑↓)各为平行和反平行折叠层，(II)表示多甘氨酸，C 为胶原蛋白。

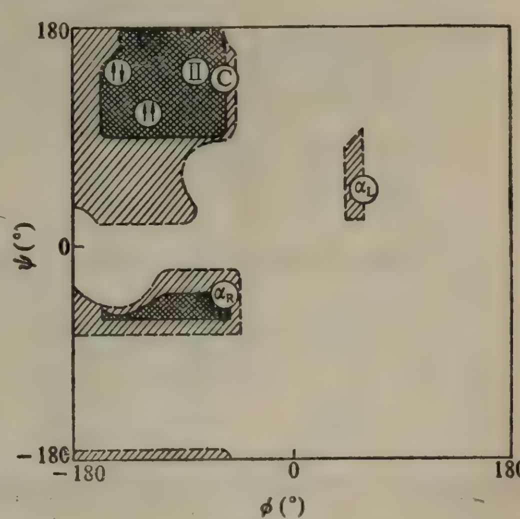


图 9-19 在多肽链中 L-丙氨酰残基的构象图

9.2.1 势能函数[46-50]

研究多肽构象起初是用硬球排斥势，以 Van der Waals 接触距离作为重叠的判别，可粗糙地估计构象空间允许的区域，如图 9-19 所示。后来应用经验势能函数来计算多肽的构象，虽然这函数还是粗糙的，而量子化学方法如半经验的方法，特别是 ab initio 能得到较精确的结果，但

对于处理生物大分子构象还是困难的，故一般仍用经验的势能方法。近年来，后者受到了人们的重视，被成功地应用到蛋白和核酸的研究上来。该方法有的称之分子力学方法。一个构象 V 的总能量是各种相互作用能的和：

$$V=V_a+V_r+V_{es}+V_{hb}+V_l+V_\tau+V_\theta \tag{9.5}$$

这里 V_a 是 Van der Waals 吸引项(London 项)，V_r 是排斥项，V_{es} 是静电项，V_{hb} 是氢键形成能量，V_l 是与键长变形有关的应变能量，V_τ 是键角变形相应的能量，V_θ 是二面角变形的能量。

非键相互作用(吸引与排斥)可取 Buckingham 6-exp 势：

$$V_a+V_r=-\frac{A}{r^6}+B\exp(-\mu r) \tag{9.6}$$

μ 取 4.6，A 和 B 为常数，相对于不同的原子对有不同的数值。

非键的相互作用也可取 Lennard-Jones 的 6-12 势：

$$V_a+V_r=-A'/r^6+B'/r^{12} \tag{9.7}$$

A' 和 B' 值表示于表 9-10 中，表中 R 为原子的 Van der Waals 半径和。

表 9-10　A' 和 B' 的数值

相互作用对	R(nm)	A'	$B'\times10^{-4}$
C…C	0.340	370	28.6
C…N	0.325	366	21.6
C…O	0.322	367	20.5
C…H	0.290	128	3.8
N…N	0.310	363	16.1
N…O	0.307	365	15.3
N…H	0.275	125	2.7
O…O	0.304	367	14.5
O…H	0.272	124	2.5
H…H	0.240	46.7	0.45

静电能：

$$V_{es}=\frac{q_iq_j}{\varepsilon r_{ij}} \tag{9.8}$$

q_i 和 q_j 是原子 i 和 j 上电荷，r_{ij} 是它们之间距离，ε 是介电常数，一般取 3.5—4，在肽单元成键原子的电荷列于表 9-11 中。

氢键能量

$$V_{hb}=A_{H\cdots X}/r^{12}_{H\cdots X}-B_{H\cdots X}/r^{10}_{H\cdots X} \tag{9.9}$$

这里 $A'_{H\cdots X}$ 和 $B_{H\cdots X}$ 是比系数。

扭转能量

$$V_\theta=\left(\frac{V_0}{Z}\right)(1\pm\cos n\theta) \tag{9.10}$$

V_0 是最高位垒

表 9-11 肽单元上原子电荷

原子	Schellman 结果	Scheraga结果	Brant结果	建议的值
C	0.425	0.449	0.394	0.4
O	−0.381	−0.416	−0.394	−0.6
N	−0.300	−0.305	−0.281	−0.3
H	0.260	0.272	0.281	0.3

应用势能函数计算丙氨酰(基)两肽的构象图表示于图 9-20。不同的研究者所用势能函数和参数不完全相同。除了应用经验的势能函数来计算肽的构象外,还有人应用半经验的和 PCILO 方法来研究氨基酸残基的构象, Pullman 夫妇[50]应用 PCILO 方法对氨基酸残基的构象做了详尽的研究。而 Hillier 和 Robson[51]还应用从头计算法计算了丙氨酰二肽的构象(见图 9-21),比较 9-19,9-20 和 9-21,它们总的是一致的,而从头计算结果进一步支持了经验势能法的结果。

上述势能函数方法和分子轨道方法同样可应用于核酸构象的计算[52-54]。

9.2.2 分子间相互作用[55-62]

在 3.2 中我们已经比较详尽地讨论过分子间相互作用。在这里,仅介绍应用势能函数方法研究分子间相互作用的问题。因为实践表明,应用势能函数法既可得到与分子轨道法(如 ab initio 计算)相符合的结果,又可克服处理稍大一些的分子所带来的困难。

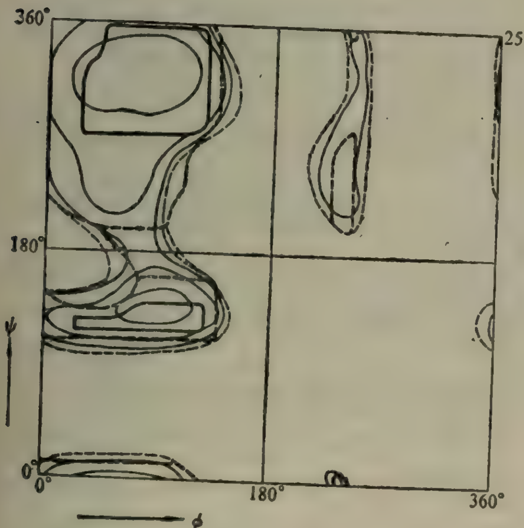


图 9-20 丙氨酰残基构象图

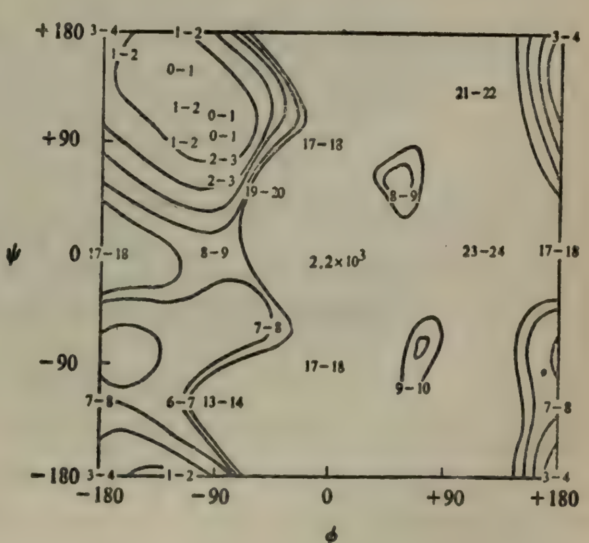


图 9-21 丙氨酰残基构象图(ab initio 法)

二分子相互作用时总的势能函数 E 可表示为：

$$E=\sum_{i\in A}\sum_{j\in B}\left\{\frac{e_i e_j e^2}{R_{ij}}-\frac{e^2}{2}(\alpha_i e_j^2+\alpha_j e_i^2)/R_{ij}^4+k_i k_j[(-A/Z_{ij}^6)\right.$$

$$\left.+(1-e_i/N_i^{\mathrm{val}})(1-e_j/N_j^{\mathrm{val}})(\exp C-\beta Z_{ij})]\right\} \tag{9.11}$$

式中 e_i,e_j 分别为分子 A 中第 i 原子和分子 B 中第 j 原子的静电荷，α_i，α_j 分别为 i 原子和 j 原子的极化率，R_{ij} 为 i 原子与 j 原子之间的距离。$Z_{ij}=R_{ij}/R_{ij}^0$，$R_{ij}^0=2\sqrt{R_i^W R_j^W}$，$R_i^W$ 和 R_j^W 分别为第 i 原子和第 j 原子的 Van der Waals 半径。k_i,k_j 是依赖于原子类别的参数。N_i^{val} 和 N_j^{val} 分别是 i 和 j 原子的价电子数。A 取 0.214 kcal/mol，C 取 4.7×10^4kcal/mol，$\beta=12.35$。如考虑氢键对总能量的贡献时，需要调整 A、B 和 β 参数[55,56]。

吴伟雄和江寿平应用上述公式编制了 ALGOL-60 语言程序，改进了求函数极小值的 BFGS 变尺度法，用于求能量极小值的构象。应用这个计算机程序计算了甲酰氨-水的相互作用[55]，碱基的水合作用[56,57]，碱基与金属离子的相互作用[58]，密码子与反密码子的识别作用等[59,60]。此外还可用 Monte-Carlo 法来模拟生物分子的水合作用[61-63]。

9.2.2.1 甲酰氨-水的相互作用

为了便于与已有的 ab initio 方法的结果比较，用改进势能法计算一个水分子与一个甲酰氨分子相互作用(图 9-22)，表 9-12 和 9-13 分别列出结构 I 及 II 一部分计算结果[55]，并与其 ab

表 9-12 结 构 I

	本工作（最优位置）	Alagona (1973)	本工作	Del Bene (1975)	Hylton (1976)	本工作	Vo Velle (1979)
θ_0	88.4	92	同左	117	90	同左	117.4
ζ_0	−2.7	−7	同左	0	0	同左	0.8
φ_0	180	180	同左	180	180	同左	180
η_0	180	180	同左	180	180	同左	180
R_0(Å)	2.74	2.815	同左	2.78	2.815	同左	2.95
E(kcal/mol)	−8.22	−9.4	−7.91	−5.06	−9.6	−8.02	−6.09

表 9-13 结 构 II

	本工作（最优位置）	Alagona (1973)	本工作	Del Bene (1975)	Hylton (1976)	本工作	Vo Velle (1979)
θ_0	96.9	108	同左	95	90	同左	107.4
ζ_0	21	12	同左	23	14	同左	11.4
φ_0	0	0	同左	0	0	同左	0
η_0	180	180	同左	180	180	同左	180
R_0(Å)	2.79	2.815	同左	2.74	2.815	同左	2.95
E(kcal/mol)	−11.08	−9.2	−9.9	−6.35	−10.8	−9.6	−7.35

initio 结果作比较(除去 Vo Velle 的结果也用势能法外)[64-67]。

从表 9-12 和 9-13 可以看出，势能函数和量子力学方法所得结果从极值位置到总能量都符合得较好，特别是最新的 Hylton 计算结果更为接近，比 Vo Velle 的结果(用势能函数法)符合的更好些。

此外，还对四个水分子与一个甲酰氨组成的"超分子"进行了计算[65]。

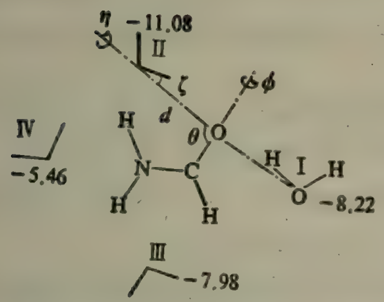


图 9-22 d 为 O—O 间的距离，θ 为C—O与O—O的夹角。把d,θ,ζ 取为变量，则总能量 $U=U(d,\theta,\zeta)$，然后求出 d_0，θ_0，ζ_0 使$U_0=U(d_0,\theta_0,\zeta_0)$取极小。

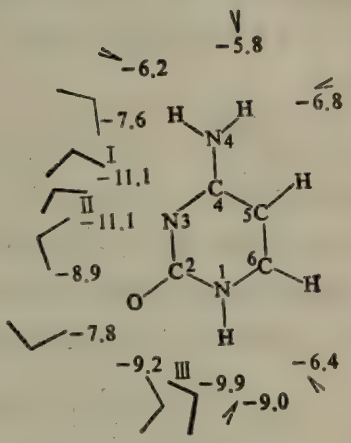


图 9-23 ab initio SCF 法计算的胞嘧啶水合作用图中粗线位置表示最佳的水合作用位置，虚线表示垂直于胞嘧啶平面且在平面之下，实线表示与胞嘧啶共面。能量单位：kcal/mol

9.2.2.2 核酸碱基分子的水合作用

对于核酸分子中碱基和水的相互作用，Pullman 等人[68,69]用 ab initio 方法做了很多工作，得到不少有益的结果。为便于与 ab initio 结果比较，这里再强调一下利用改进的分子间相互作用势函数对胞嘧啶和腺嘌呤等与水的相互作用所作的计算[56](参见 4.7)。

1. 胞嘧啶的水合作用

在相互作用时，胞嘧啶和水分子都可以在三维空间内任意运动。计算得到的最优水合位置和对应能量如图 4-35 所示。为了与之比较，图 9-23 给出了 Pullman 等[68,69]人用 ab initio 法计算所得到的结果。由图 4-35 和图 9-23 可以看出，改进势能法和精确的 ab initio 法的结果是接近的。

胞嘧啶水合作用的最优位置说明，胞嘧啶的最优水合位置分布在二个不同的区域，即胞嘧啶的 N 3 原子和 N 1 原子的邻近区域，而在其它区域，它的水合作用能没有与之可以比较的极小值。由图 4-35 可知，当水分子与胞嘧啶分子共面且水分子中一个H原子和胞嘧啶的N 3 原子形成氢键时，其相互作用能有极小值(−11.30 kcal/mol，−11.67 kcal/mol)，其中极值位置 I 也是其相互作用能最小值位置，该 I、II 位置就是胞嘧啶的第一第二最优水合作用位置。胞嘧啶的另一个水合作用的最优位置是在它的C2＝O 基团和 N 1 H 基团之间，在该位置时水分子也与胞嘧

啶分子共面。这时相互作用能的贡献主要来自于胞嘧啶的 N 1 H 基团和C2＝O 基团中O原子与水的相互作用。总之图 9-23 简单地反映了胞嘧啶水合作用的图象。

2．腺嘌呤的水合作用

计算方法同上，结果见图 4-35 和图 9-24。它表明腺嘌呤的水合作用有三个较强的区域：

(1) 当水分子处于图 4-35 中极值位置 I 时其相互作用能有最小值(－13.67 kcal/mol)，位置 I 即为腺嘌呤的第一最优水合作用位置。此时水分子不与腺嘌呤分子共面，它与腺嘌呤的 N 6 H、N 7 原子形成桥式连接，其中水分子中的O和一个H原子分别与腺嘌呤的 N 6 H 和 N 7 原子形成氢键。

(2) 水合作用的第二个最优位置是在区域 N3—C4—N9H 附近，此时水分子也不与腺嘌呤共面，它与腺嘌呤的 N 3，N 9 H 形成桥式连接，此时相互作用能的主要贡献来自 N 9 H 和 N 3 原子与水的相互作用，其相互作用能为－11.30 kcal/mol。

(3) 当水分子处于图 4-35 中极值位置 III 时其相互作用能有极小值（－8.02 kcal/mol)，此时水分子与腺嘌呤分子共面且水分子的一个H原子与 N 1 原子形成氢键。

由图 4-35 和图 9-24 的比较可知由改进势能法计算的结果是接近于 A. Pullman 等人的 ab initio SCF 法计算结果的。

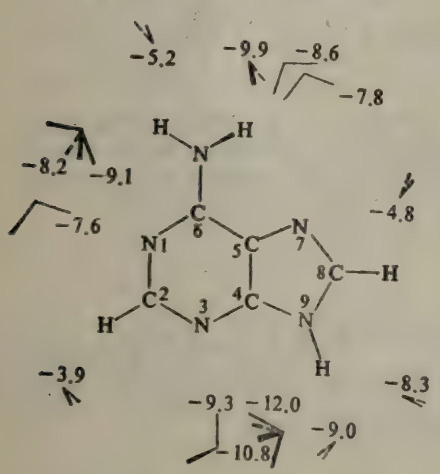


图 9-24 ab initio SCF 法计算腺嘌呤的水合作用 粗线位置表示其最佳水合作用位置，实线：表示水分子与腺嘌呤共面。虚线：表示水分子垂直于腺嘌呤平面且在平面之下。能量单位：kcal/mol

由上述的胞嘧啶、腺嘌呤的水合作用计算结果也反映了它们的水合作用最优位置和水合作用能的分布是与碱基分子中各原子的净电荷分布紧密相连的。在胞嘧啶中因 N 3 原子的净电荷最负(－0.663) 而胞嘧啶水合作用的第一最佳位置就是在 N 3 附近的最小值位置 I。在腺嘌呤中带负电荷的各原子的净电荷按其绝对值大小排列的顺序是：N 7（－0.557)，N 1（－0.518)，N 3（－0.510)，N 6(－0.439)，N 9（－0.192)。而用势能法得到的腺嘌呤水合作用的最佳位置的排列顺序是与之对应的，即N 7 —N 6，N 3 —N 9 H，N 1 区域。由此可见由改进势能法计算的碱基的水合作用能是与它们的原子净电荷分布紧密相连的。

此外还利用此势能法计算了 B-DNA 分子中互补碱基对 G-C、A-T 水合作用[58]，除了应用势能法可以用来研究生物分子水合作用，应用蒙特卡罗模拟也是一个很好的方法[61,62]。

9.2.2.3 阿霉素与核酸的相互作用

李光等应用分子之间相互作用势能法和自编的微机程序研究了抗癌药物阿霉素和核酸的嵌插作用。他们的计算结果和他人的理论计算和实验结果列于表 9-14 中。

他们取 ADM 三个自由度变化：即 Δx，Δy和 $\Delta \phi$，它们都是相对初始位点而言的变化量，$\Delta \phi$ 为 ADM 绕通过自身四环上 M 点的 z 向轴的转动量。计算结果，在小沟嵌入方式中找到

表 9-14 ADM构象计算结果与他人工作比较

研究方法	化合物名称	ϕ_1	ϕ_2	ΔE(kcal/mol)	研究者
X光衍射	Daunomycin(DNM)	292°	125°		Neidle
	Carminomycin	290°	117°		Von Dreele
	N-Bromoacetyl DNM	281°	102°		Anguili
	DNM-DNA(66p)	282°	91°		Quigley
理论计算	DNM	290°	93°	0	Neidle
	DNM	292°	96°	0	Brown
	DNM (a)	288°	79°	18	Nakata
	(b)	258°	259°	0	
	ADM (a)	288°	77°	1.5	Nakata
	(b)	258°	257°	0	
	ADM	288°	103°	0	温元凯等

了一个稳定点(S)；在大沟嵌入方式中找到了两个稳定点(V)和(P)，(图 9-25)。能量最稳定点为(P)。

9.2.3 核苷酸和蛋白质空间构象和电子结构的计算

9.2.3.1 经验势能法

许多实验事实证明，蛋白质空间结构信息是包含在它的氨基酸顺序中的。可以认为，在一定环境中天然蛋白质分子是处于全系统自由能最小的状态，因而使我们有可能从它的一级结构来预测它的高级结构。虽然这样的预测还存在不少困难，但已经得到一些有益的结果，它有助于了解蛋白质的折叠，能合理估计不同结构的稳定性，不仅有理论意义而且有实际意义[70,71]。

在现有的各种预测三级结构的方法中。我们仅讨论势能方法。对于不大的蛋白质分子，其可能的构象也达到惊人的数字，如一个 150 个残基的多肽链，可能的构象有 10^{45}种[71]。如果计算的话，则需很多计算时间，为了减少庞大的计算量，需要对多肽链作些简化，Levitt[73]提出了一种简化(如图9-26所示)。这个简化模型是把一个多肽链主链以一个虚键来代表，每一个C^{α}之间距离是固定的，围绕每一个虚键以一个 α 角取代原来的两个 ϕ 和 ψ 角。Levitt 等基于这个简化方法较好地模拟了牛胰蛋白酶抑制剂(BPTI)，除此之外，还有其它一些方法。Scheraga和他的同事于1975年编写了多肽的经验构象能程序(ECEPP)，1982年又发表了进一步改进的ECEPP/2 程序[73,74]。Levitt 等编成类似的程序[75]。1985年 Fraga 编写预测蛋白质三级结构的程序[76]。

1. 多肽三级结构预测程序[77]

通过计算分子的势能来寻找具有最低势能的分子的三级结构包括两个方面的问题：

(1) 选取适当的能量计算公式，它能表示为分子中各结构参变量的函数，且函数形式要简单，便于计算。

(2) 选取适当的最小化方法，使得函数能在尽可能少的时间内收敛至最小值。

这样，寻找具有最小势能的分子结构归纳为数学问题便是：在蛋白质的结构空间中寻找一

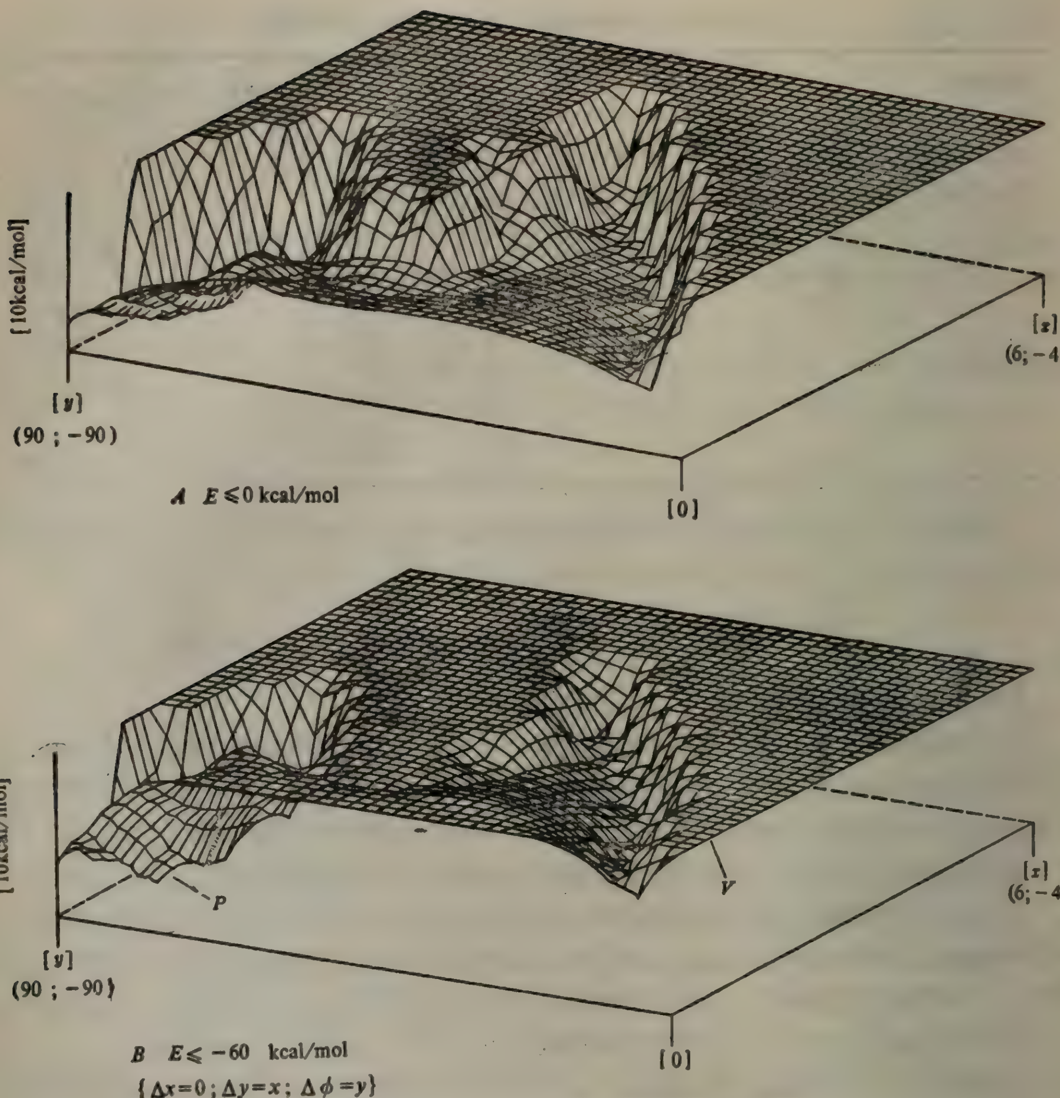


图 9-25 ADM-d(C-GC)₂ 大沟嵌入的两种势能曲面图

组适当的结构参数，使得势能函数具有最小值。

这里采用了 Scherga 的经验势能函数：$E=E_{ES}+E_{NB}+E_{TOR}+E_{LOOP}+E_{CYSTR}$，其中

① 静电能：

$$E_{ES}=\sum_{ij}332.0\times q_i\times q_j/D\times r_{ij} \tag{9.12}$$

q_i,q_j 分别为二相互作用的 i 原子、j 原子的电荷（电子电荷单位），r_{ij} 是它们的核间距(Å)，$D=2$ 为介电常数。

② 非键相互作用能(包括氢键相互作用能)：

$$E_{NB}=\sum_{ij}(FA^{KL}/r_{ij}^{12}-C^{KL}/r_{ij}^{6}), \tag{9.13}$$

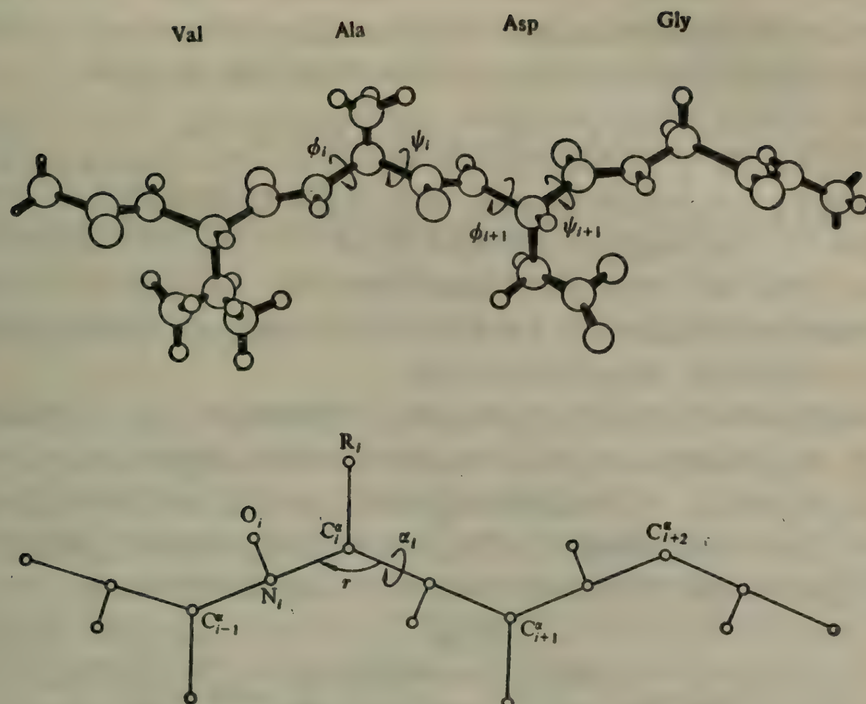


图 9-26　多肽链的简化表示

其中系数 FA^{KL}、C^{KL} 对于特定类型(K和L)的二原子相互作用为一常数。r_{ij} 的意义同 a)。

③ 扭转能：

$$E_{TOR}=\sum_k (u_0/2)\times(1\pm\cos n\theta_K),\tag{9.14}$$

$\sum_k$ 对所有的肽键二面角 ω，侧链二面角 χ_i 及末端基团的二面角(若有的话)求和。u_0、n 对于特定的 θ_k 为一常数，本文中取 $\omega=180°$。

④ 二硫键成键能量：

$$E_{LOOP}=\sum_k B[(r_1-r_{10})^2+(r_2-r_{20})^2+(r_3-r_{30})^2],\tag{9.15}$$

$\sum_k$ 对分子中所有可能的二硫键求和，假设 ith 残其与 $j$th残基形成一胱氨酸。用图 9-27 表示二硫键的形成，则 r_1、r_2、r_3 的意义分别为 S_iS_j、$C_i^\beta S_j$、$S_iC_j^\beta$ 的距离，r_{10}、r_{20}、r_{30}分别为标准二硫键中相应的距离，B 为任意选择的常数。

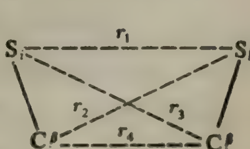


图 9-27

⑤ 二硫键扭转能：

$$E_{CYSTR}=\sum_k A(r_4-r_{40})^2,\tag{9.16}$$

$\sum_k$ 对分子中所有可能的二硫键求和，r_4 为$C_i^\beta C_j^\beta$ 间的距离。r_4 为标准二硫键中相应的距离。

为了有效地寻找到具有最低势能的分子的三级结构，还须选择一套有效的能量最小化方法。在此，我们选择了变尺度法和变量轮换法。变尺度法是近年来被证明收敛快、稳定性都较好的方法，也是人们常用的方法。然而，变尺度法需要计算能量梯度，按照下式：

$$\frac{\partial E}{\partial \theta_m}=\sum_{i}^{N}\sum_{j>i}\frac{\partial E}{\partial r_{ij}}\cdot\frac{\partial r_{ij}}{\partial \theta_m},m=1,2\cdots\cdots M, \tag{9.17}$$

其中 M 为变量个数，N 为原子个数的总和。可知，计算梯度所需的时间为 $O(N^2M/2)$。当变量个数增加时，$O(N^2M/2)$远远大于计算能量所需时间 $O(N^2/2)$。此外，由于变尺度法收敛与起始点很有关系，往往只能求出局限于起始区域的极小值，故我们还采用了变量轮换法，它只需计算能量值不需计算梯度值。仅当变量轮换收敛到一定值时，才开始用变尺度法。变量轮换法的采用对于改进上述二点不足有一定的作用。

刘志平和江寿平用 FORTRAN IV 语言编制了一套预测程序，它由一组具有特定功能的子程序组成，输入多肽数据，输出多肽分子的具有最低势能的三级结构(原子的笛卡儿坐标或在任一方向上的投影图)。关于各程序的功能及它们之间的调用，我们将结合图 9-28 讨论。

图 9-28 中虚线表示程序或数据块的调用，箭头所指为被调用的程序，线上方的注解为程序间数据传送的内容(包括被调用的程序的输出结果)。箭头所指为传输方向。

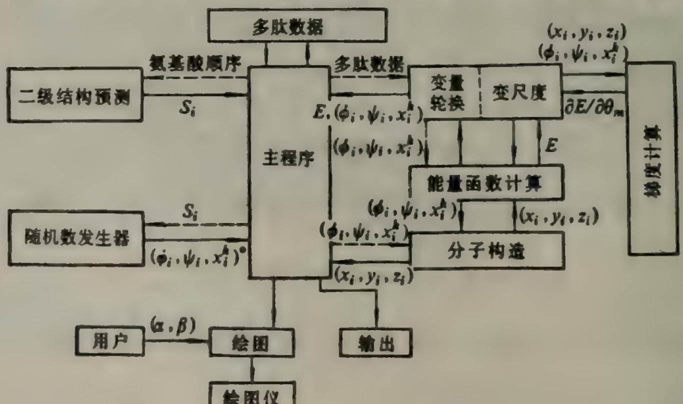


图 9-28

(1) 主程序完成多肽数据的输入，调用各子程序，输出结果等功能。多肽数据包括：(1) 多肽的氨基酸顺序；(2) 每一个氨基酸残基的立体化学标志(D型/L型)；(3) 选择输入二硫键的配对情况及它们的个数；(4) 希望优化的变量名称。它可由用户组织成盘文件进行输入或通过键盘输入。此外，主程序还读入氨基酸残基的标准数据。

(2) 二级结构预测子程序。此程序输出的各残基的二级结构值(S_i)可部分解决起始构象选择的盲目性，减少优化过程所需的计算时间。处于不同的二级结构状态的氨基酸残基有不同的(ϕ_i, ψ_i, χ_i^k)值，它可导致不同的起始构象，对于α-螺旋残基，我们在区域(-65°, -55°)之间随机选取 ϕ_i，在区间(-45°, -35°)间选取 ψ_i，对于 β-折叠残基，ϕ_i 取(-130°, -110°)间的值，ψ_i 取(140°, 160°)间的值，其它残基 ϕ_i 和 ψ_i 分别在(-120°, -60°)及(60°, 120°)间取值。对所有残基的 χ_i^1 取60°，$\chi_i^k(k>1)$取180°。

(3) 随机数发生器：利用随机函数在上述范围内随机产生(ϕ_i，ψ_i)值，不同的取值可产生不同的起始构象，经多次取值后，选择能量最低的那组参数作为起始构象参数(ϕ_i，ψ_i，χ_i^k)^0。

(4) 能量优化子程序：从起始构象开始，经最小化过程得到一组具有最低势能的构象参数。这个子程序由二部分组成，变量轮换能量最小化，每次按顺序依次改变一个变量，若当前势能值小于上一次计算的势能值，则用当前构象参数值取代上次计算所用参数，否则不变。

能量优化子程序调用梯度计算子程序和能量计算子程序，后者输出能量梯度值$\frac{\partial E}{\partial \theta_m}$和能量值 E。

(5) 分子构造子程序：利用参数(ϕ_i，ψ_i，χ_i^k)计算分子中各原子的笛卡儿坐标(x_i，y_i，z_i)。

(6) 绘图子程序：主程序调用绘图子程序，通过绘图程序在x-y 绘图仪上输出多肽分子的三级结构在任一方向上的投影图。用户可通过人机对话，告诉计算机投影方向(α，β)。

(7) 输出程序输出原子的(x_i，y_i，z_i)坐标和对应的能量值 E。

2．乙型肝炎病毒的表面抗原蛋白分子部分片段的三级结构[77,78]。

前面已讨论了 HBsAg 的 4 种亚型(adr、adw、ayw和adyw)的二级结构研究，并指出，HBsAg 的抗原专一性不仅取决于抗原决定簇的性质和数目，而且还决定于它们的空间分布。4 种亚型的二级结构表明，抗原决定簇所在区域以及它附近的区域(120—150)的二级结构虽有一定特征，但都很相似。实验结果表明，这段区域在不同亚型之间存在着抗原性强弱的差异，即便是同一亚型，不同片段其活性也是不一样的。因此，要进一步了解 HBsAg 抗原决定簇的抗原性的分子机制就有必要研究它们的三级结构。

采用了 Scheraga[73]的经验势能公式，结合变量轮换法和变尺度能量最小化方法，通过最小势能函数，寻找到了 HBsAg 亚型 adw、adr 和 ayw 的片段 122—132，亚型 adr 的片段 122—137、125—137 和 128—137 的具有最低势能的三级结构[77]。

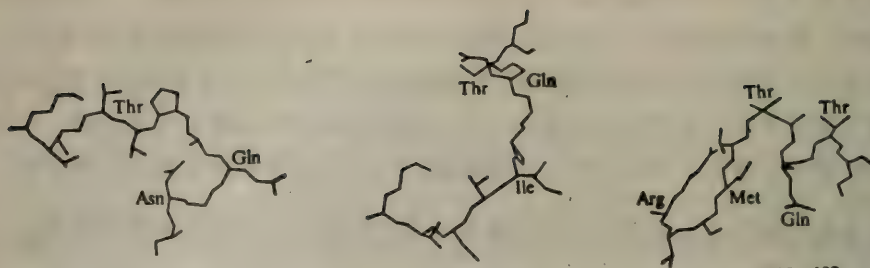
图 9-29、图 9-30分别给出了亚型 adr、adw、ayw 中片段 122—132 及亚型 adr 中片段122—137，125—137及128—137的三级结构预测结果，我们分两方面讨论这些片段的结构异同：

(1) 亚型 adw、adr、ayw 中片段 122—132 的三级结构比较：图 9-29 下方的三段顺序分别为亚型 adw、adr及ayw区域 122—132的一级结构。从图可以看出它们在一级结构上的差异。二级结构研究表明：这段区域的二级结构在三种亚型之间很相似，无太大差异，然而三级结构研究表明它们在三级结构上存在着明显差异。adw中，第 126 位、131位残基分别为 Thr (苏氨酸)和 Asn(门冬酰胺)，Thr(126)-Pro-Ala-Gln-Gly-Asn(131)形成一环结构。

adr 中，由于第 126、131 二位上氨基酸残基的替换(Thr 换为 Ile，Asn 换为 Thr)，原存在于 adw 中 Thr(126)和 Asn(131)间的氢键消失，新的氢键在 Gln(129)与 Thr(131)之间形成。

ayw 122—132 的三级结构与前两种亚型的差异很大，在 128—131 处有一β 转角形成。

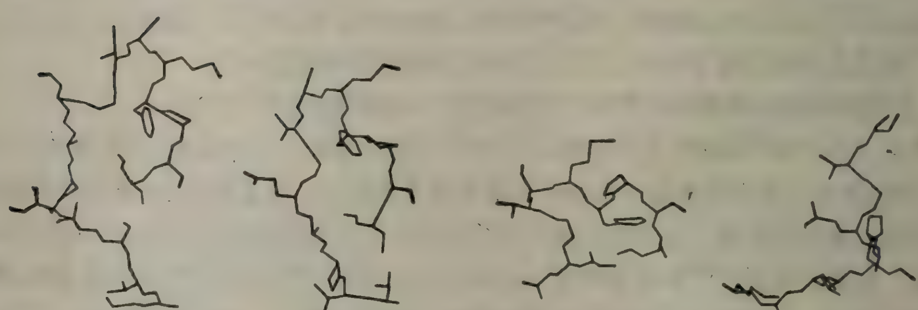
(2) 亚型 adr 122—137、125—137 和 128—137 片段的三级结构比较：从图 9-30 A，B 可以看出，122—137 片段的三级结构中，区域 125—137 处的结构片段 125—127 基本一致，在 124 及 127 二氨基酸残基之间有个二硫键形成。片段 128—137 的三级结构(图 9-30 C)与前二者相



adw122－132　　adr122－132　　ayw122－132

A　　*B*　　*C*

图 9-29　*A*.adw 122-Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn Ser
B.adr 122-　　Ile　　Thr
C.ryw 122-Arg　　Met Thr Thr



adr 122－137 (Cys124－137)　　adr125－137　　adr 128－137　　adr122－137

A　　*B*　　*C*　　*D*

图 9-30　adr 122-Lys Thr Cys Thr Ile Pro Ala Gln Gly Thr Ser
Met Phe Pro Ser Cys

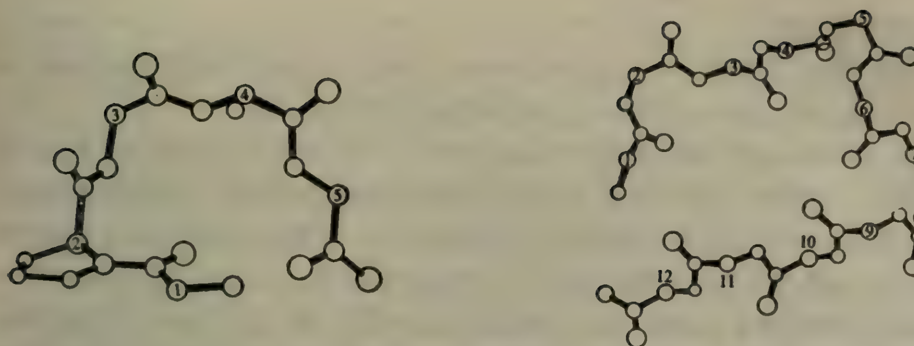
差较大。图 9-30 *D*是 adr 122—132 片段的三级结构在另一方向上的投影图。

3．胰岛素原 C-肽诸片段的构象分析[79]

在胰岛素原的生物合成中，C-肽对于促进胰岛素的二三级结构的形成起着很大作用，它还可作为胰岛 β-细胞的方便可靠的探针而广泛应用于临床。

在人工合成 C-肽时，发现除羧端肽 C15—19 与氨端肽 C 20—31 的接肽反应 产 率仅理论值的80％外，其它的均接近理论计算值。人们不禁要问接肽反应的二个肽段本身 的 构象对于其接肽反应产率是否有关系？我 们计算分析了 4 步接肽反应中 8 个肽段的构象，发现 两 个肽段的构象确 实与接肽反应的产率有联系。

肽 C15—19 及 C20—31 的三级结构透视图如图 9-31。从此图可见，这两肽都具有发夹结构，他们的 N 端靠近其C端，可形成氢键。这种结构特征在其它接肽反应中氨端肽及羧端肽中的构象中并不存在。由于 C15—19 和 C20—31这两段肽存在发夹结构，它 们的 C 端与 N 端靠近又形成氢键，使得 C20—31 的 N 端与 C15—C19 的 C 端接触变得困难，从而减少它们的碰撞机会，因而使得这两个肽的接肽反应率较之其它接肽反应产率比理论计算值低。

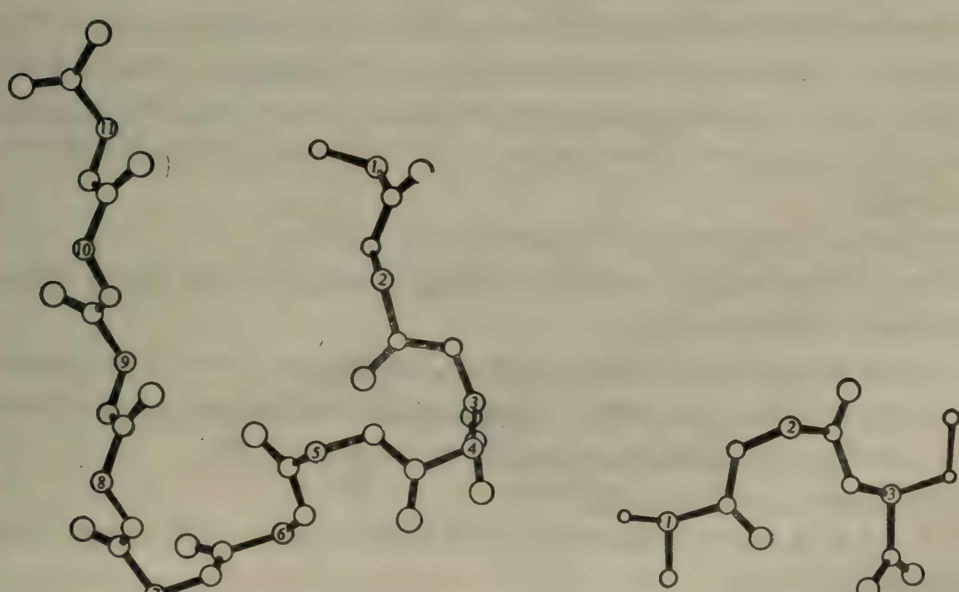


H—Gly·Pro·Gly·Ala·Gly—OH[C15—19]　　H—Ser·Leu·Gln·Pro·Leu·Ala·Leu·Glu·Gly·Ser·Leu·Gln—OH [C20—31]

图 9-31

因为蛋白质的三级结构是由其一级结构决定的，因此，可以改变合成路径，即选择合适的片段，使它们的构象有利于接肽反应的进行，以此来提高接肽反应产率。改变了新的策略合成C肽，后一策略的C15—31的产率的理论值提高到90%，我们对新路径的肽段进行了计算，图9-32是C18—20和C21—31的构象图。它的构象与前接肽反应中C15—19和C20—31的构象完全不同，它们的N端与C端都分的很开。

由上可见，在多肽合成前路线设计，可预测它们的构象而为设计合成路线提供有益的信息，



H—Leu·Gln·Pro·Leu·Ala·Leu·Glu·Gly·Ser·Leu·Gln—OH　　H—Ala·Gly·Ser—OH[C18—20]

图 9-32

减少盲目性。

9.2.3.2 分子轨道方法

量子化学计算中长期以来未得到令人满意解决的两个问题，即提高计算速度和确保叠代收敛，对大一些的分子，尤其是生物分子的研究，这两个问题尤其突出和重要。

Hartree-Fock-Roothaan(HFR)方程是 ab initio，INDO，CNDO等分子轨道计算方法的基础，由于 HFR 方程是一个所谓“自递归定义”的方程，只能用叠代的方法求解。但简单叠代法常常发生叠代发散的现象。对于含过渡元素的分子，极少收敛，分子大到一定程度时，几乎都不收敛，就是一些简单分子也出现了叠代发散情况[80-82]。普遍应用的 Gaussian 76 报告指出[83]，不收敛仍是自洽场叠代中最为常见的困难。由于不能克服叠代发散问题，在生物体中至关重要的 ATP 分子的 CNDO 研究也未能进行[2,84-88]。为了解决自洽场叠代中收敛问题，已作了很多研究，提出了不少方法，但这些方法只能对某些分子有效。

为了解决自洽叠代中收敛问题，我们提出了一个新的叠代公式，称之为对角位移公式[89]（参见 4.2)。

量子化学计算，就是半经验的 CNDO 方法，计算量和存贮量都很大，从头计算所需的计算量就更加庞大。国际上流行的 Gauss 系列程序，一般只能计算到 80 STO 基函数至多不超过 120 个[90,91]。由于目前所使用的各种类型的从头计算程序，都是以 Gauss 函数来进行排斥积分计算的，它是极为耗时的。因而如何提高排斥积分计算速度成了一个很有实际意义的工作。由于在为数庞大的排斥积分中有相当大一部分数值是很小的，但它在积分总数中占很大比重，而它们对计算结果影响很小，分子越大，这效果越明显。为了得到 Gauss 函数积分的估计式，以加速计算速度，我们提出了一个排斥积分估计公式，大大加速了积分计算[92]。宣建成和江寿平应用上述的“对角位移”公式和排斥积分估计公式用 Fortran 编成了大型机和微机上运行的 CNDO 和 ab initio 程序一套。能在 IBM-PC（256 K 内）进行 80 个 STO 基的从头计算和 126 个 STO 基的 CNDO 计算。能在 1 M 以上内存的小、中型计算机上进行 250—400 个 STO 基的从头计算或 500 个以上的 STO 基的 CNDO 计算工作。由于采用了“对角位移公式”，因而任意分子均可收敛，计算例见后。

由于大分子从头计算法耗时甚多，下面均以 CNDO 计算为例。

(1) 环化鸟苷酸双聚体有 68 个原子 216 个轨道。用简单叠代法发散，用对角位移公式仅叠代 6 次即达到 10^{-5}精度(表 9-15)。

(2) 铁叶啉双聚体有 406 个轨道。是目前国内文献报道中用 CNDO 计算的最大分子。由于有 2 个过渡元素 Fe 原子，简单叠代法一开始就发散，用对角位移法叠代10次即达到 10^{-5}精度(表 9-16)。由叠代结果可以看到其叠代能量为严格下降(除第二次外)。

1. 腺苷三磷酸(ATP)的从头计算[93-95]

在 4.2.5.2 中，作为例子，我们已经讨论过腺苷三磷酸(ATP)分子的电子结构。在这里，要补充讨论的是：

(1) AMP、ADP 和 ATP 分子的电荷分布，在腺嘌呤部分的电荷用不同方法得到的结果大致趋势相似(原子上电荷的正或负)。

表 9-15 $(cGMP)_2$计算结果

叠代编号	电子能量 (a. u.)
1	−3318.798
2	−3319.282
3	−3319.538
4	−3319.800
5	−3319.855
6	−3319.878

表 9-16 $O(FeTPP)_2$ 叠代情况

叠代编号	电子能量 (a. u.)
1	−8503.429
2	−8445.981
3	−8490.683
4	−8495.635
5	−8501.923
6	−8506.087
7	−8506.614
8	−8507.084
9	−8507.223
10	−8507.286

(2) 由图 9-33 可见，磷酸基团上氧原子带有相当大和几乎相等的负电荷。因此，它们和金属离子相互作用的可能性几乎相等，这在生物学上有着重要作用(参见第 10 章)。

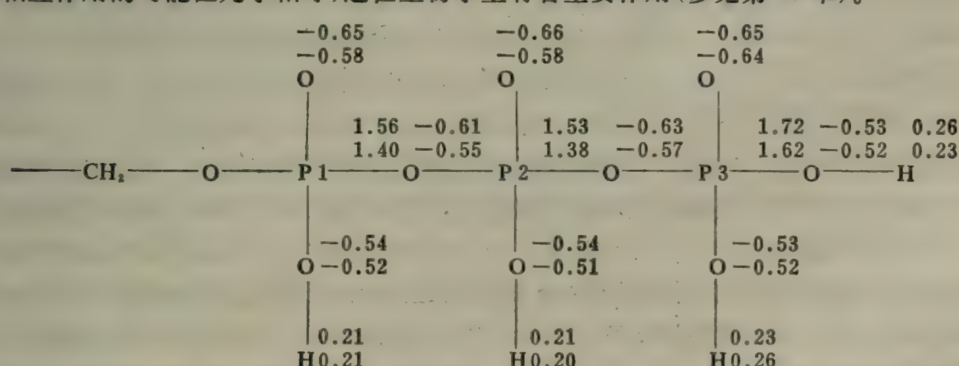


图 9-33

(3) 所谓"高能磷酸键"是指水解自由能，而不是能量贮存在 P—O 键，再次强调这一点，对于生物学研究同样是重要的。

对于腺苷三磷酸(ATP)分子的分子轨道研究，在方法学上，经历了简单的 Hückel方法、EH-

MO 法、CNDO 法和现在使用的 ab initio 法，为我们研究 ATP、AMP 和 ADP 的电子结构和生物学功能提供了更多的手段和有益的信息。

2．环化鸟苷酸和环化腺苷酸的从头计算[96,97]

环化鸟苷酸(cGMP 或 cyclic GMP)和环化腺苷酸(cAMP 或 cyclic AMP)一样，作为传递激素和其它刺激的第二信使，它们都具有重要的生物学功能，也是生命过程中的重要分子。

江寿平等曾用 EHMO 法、CNDO 法和 ab initio 法对这 2 个分子进行了理论计算[84,96-98]，研究了它们的电子结构、构象和作用机理，从分子和电子水平上探讨了它们与激酶等的相互作用。

我们知道，cAMP 和 cGMP 是由碱基和核糖与磷酸双环两部分组成的，而碱基可以相对于核糖-磷酸部分绕糖苷键转动造成顺式和反式构象，这 2 个分子的结构示于图 9-34 和图 9-35。

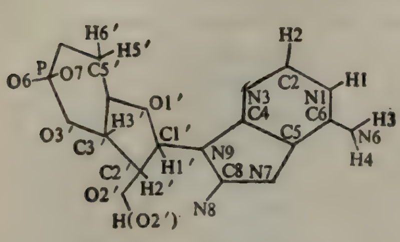


图 9-34 cAMP 分子的结构

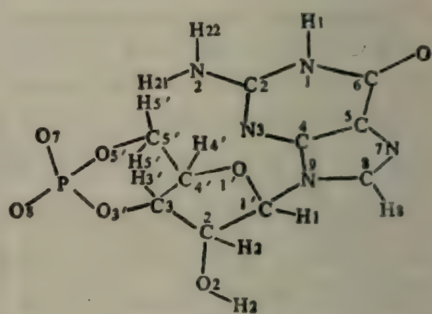


图 9-35 cGMP 分子

(1) cAMP 的计算结果

用从头计算法计算cAMP分子，得到它的总能量是－1353.27539 原子单位，它的偶极矩是33.1562 德拜。而用 CNDO 方法得到的分子总能量是－255.08594 原子单位，偶极矩35.1520德拜。两种方法所得到的偶极矩十分近似，而 Jordan 用 EHMO 方法所得到的总能量和偶极矩分别为－55000 kcal/mol和2 7.20 德拜。这分子的原子净电荷的计算结果列于表9-17中。从表9-17中可见，用 EHMO，CNDO 和从头计算法所得到的原子净电荷其大致趋势是一致的(电荷的正或负)， 但在数值上是很不一致， CNDO 的结果与从头计算法的结果比较接近，而 EHMO 的原子净电荷结果是夸大了。在这分子的碱基和核糖部分上最负的原子是不一致的。在核糖和碱基上，用从头计算法得到的这两部分上最负的原子各为 O 3′和 N6 ，而用 EHMO 得到的是 O1′和 N7 .而 CNDO 的结果和从头计算法结果一致，对腺嘌呤分子，Pullman 和 Clementi 等人都用从头计算法进行了计算，他们所得到的净电荷数值相差也较大，但是最负的原子都是 N6 ，和这里的计算结果一致。核糖上的氢原子，三种方法都得到 H(O 2′)原子上净正电荷最大，而有趣的是它比其它的 H 原子大好几倍，因而它是最可能成为氢键的 H 供者，实验也证实它在 cAMP 和蛋白激酶相互作用中起作用。在磷酸部分，P 原子的净电荷数值，因在我们计算中没计及 3d 轨道，因而偏大，但不影响我们这里的讨论。 和磷原子相连的两个氧原子，用从头计算法和 CNDO 方法得到它们是 cAMP 上最负的原子， 实验证明它们也是这分子上最活泼最敏感的原子。它们在和其它分子相互作用中将起重要作用。根据这分子的电子结构的特点，下面将讨论

cAMP 和蛋白激酶的相互作用。

凡有 cAMP 的细胞,都有一类能催化蛋白质产生磷酸化反应的酶——蛋白激酶。cAMP 通过蛋白激酶发挥它的作用,当它结合到无活性的蛋白激酶调节亚基上时,使无活性的催化亚基和调节亚基复合体解离,产生了有活性的催化亚基和 cAMP-调节亚基复合体,对于 cAMP 和调节亚基如何相互作用,用化学修饰法做了很多工作。但可以从 cAMP 分子的电子结构来提供分子相互作用的有用信息。可以设想,和磷原子相连的两个氧原子,由于带有较负电荷,它与蛋白激酶调节亚基上带正电荷的部位产生静电相互作用,这个远程力可能在识别和定向方面起作用,而使两分子相互接近,接着短程力起作用,酶分子产生构象变化,使调节亚基和 cAMP 相互结合[98]。碱基处于激酶调节亚基的疏水区域,带较大负电荷的 O 3′和 O 5′以及带正电荷的 H(O 2′)和调节亚基上的氨基酸残基侧链上氢和氧等原子形成氢键,因而可以得到 cAMP 。

cAMP 受体蛋白(CAP)是一种变构蛋白,它在 cAMP 存在时才和 DNA 专一的结合。X 光衍射表明了 cAMP 是完全埋在 CAP 的二聚体内部,cAMP 上磷酸基团上两个氧原子带有大而几乎相等的负电荷,它们可和 CAP 上第 82 个精氨酸残基侧链上两 NH_2 上的正电荷产生静电作用。而 O 5′和第 83 个丝氨酸残基侧链上 OH 形成氢键。在碱基上,除 N1 原子外,N7 带有较负电荷,而和 N1 相连的两个氢原子带有正电荷,在 cAMP 处于反式构象时,它们可以与 CAP 上相近的残基上的氢和氧原子分别形成氢键,根据 cAMP 的电子结构和 cAMP-CAP 复合物的 X 光衍射结果可得到它们的部分结合部位(图9-36)。

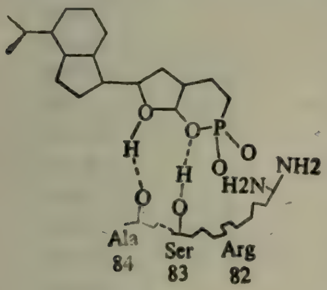


图 9-36

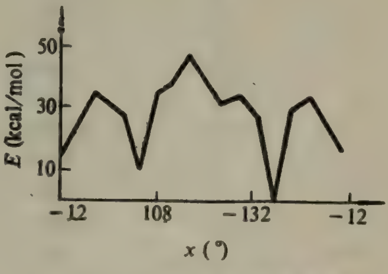


图 9-37 cGMP 分子的能量曲线

从上述可见,由分子的电子结构结合某些实验资料可以对生物分子的作用机理更好地了解,它可以更好地说明某些实验结果,更进一步说明结构与功能的关系。

(2) cGMP 的计算结果

江寿平和宣建成用从头计算法和 CNDO 方法得到的 cGMP 分子的原子净电荷列于表9-18中,为比较方便,表9-18中还列出 Clementi[3]等人的单独鸟嘌呤分子的从头计算结果。从表中可见 CNDO 的结果与从头计算的结果大致相近,在碱基上最负的原子,从头计算结果是 N2,其次是 N1,但是 CNDO的结果是 N 3,其次是 N 2。在核糖部分,较负的原子是 O 3′和O 5′,其次是 O 2′,而 H(O 2′)原子带有很大的正电荷。在磷酸部分,O 7和O 8带有较负电荷,这部分也是较活跃部分,这些电荷分布特点,在生化反应中将起重要作用。

cGMP 分子也可以绕糖苷键C1′—N9转动,粗略地说,当碱基位于接近核糖位置时,是顺式

表 9-17

方　　法	ab initio	CNDO	EHMO[843]
C1′	0.1394	0.2072	0.567
O1′	−0.1849	−0.2648	−0.854
C2′	0.0559	0.1107	0.333
O2′	−0.2128	−0.2362	−0.806
C3′	0.0306	0.1323	0.377
O3′	−0.3593	−0.4321	−0.435
C4′	0.0548	0.1210	0.396
C5′	−0.1110	0.0977	−0.038
C6′	−0.3158		−0.166
O5′		−0.4476	
O6	−0.6402	−0.6756	−0.638
O7	−0.6727	−0.7216	−0.627
P	1.0823	1.3392	0.093
H1′	0.0443	−0.0157	0.054
H2′	0.0302	−0.0196	0.055
H3′	0.0319	−0.0348	0.054
H4′	0.0402	−0.0051	0.047
H5′	0.0502	−0.0111	0.085
H6′	0.0289	−0.0265	0.062
H7	0.0386		0.147
H8	0.0131		0.129
H(O2′)	0.1381	0.1278	0.418
N1	−0.2844	−0.1145	−0.080
C2	0.1947	0.2262	0.511
N3	−0.1416	−0.1355	−0.773
C4	0.2073	0.2265	0.465
C5	0.0150	−0.0426	0.096
C6	0.5393	0.3211	0.575
N6	−0.3857	−0.2170	−0.349
N7	−0.2141	−0.1830	−0.834
C8	0.1717	0.1874	0.463
N9	−0.1749	−0.0840	−0.086
H1	0.2626	0.1642	0.238
H2	0.1185	0.0510	0.044
H3	0.2410	0.1546	0.234
H4	0.2667	0.1767	0.235
H8	0.0841	0.0271	0.234

构象，而当碱基远离核糖时，称之反式构象，它的构象的计算结果表示于图9-37中，这分子最稳定的构象是分子总能量最小时的构象。由图9-37可见，它有 3 个能量谷，其最小的位于－102°处。这与 X光衍射结果一致。

实验证明，cGMP 像 cAMP 一样，也能活化依赖于 cAMP 的蛋白激酶。由图9-37可见，它的位置很高，故它与蛋白激酶相互作用时，可以认为 cGMP 还是处于原来顺式的构象。

表 9-18 cGMP 原子净电荷

	ab initio	CNDO	ab initio[3]
N1	−0.3722	−0.2222	−0.4720
C2	0.3879	0.4005	0.3919
N2	−0.4376	−0.2547	−0.6135
N3	−0.3190	−0.2694	−0.3611
C4	0.1676	0.2101	0.1929
C5	−0.0336	−0.0467	−0.0497
C6	0.2545	0.3383	0.3590
O6	−0.2106	−0.2852	−0.3684
N7	−0.2057	−0.1469	−0.2392
C8	0.1013	0.1645	−0.0015
N9	−0.2064	−0.1118	−0.4470
H1	0.2194	0.1725	−0.3519
H2	0.2217	0.1822	0.3489
H3	0.2207	0.1706	0.3133
H8	0.0529	0.0342	0.2133
C1	0.1342	0.2145	
O1	−0.1702	−0.2348	
C2	0.0265	0.1036	
O2	−0.1991	−0.2371	
C3	0.0135	0.1343	
O3	−0.3456	−0.3766	
C4	0.0054	0.1031	
C5	0.0202	0.1383	
O5	−0.3672	−0.3810	
P	1.2921	1.5341	
O7	−0.3049	−0.2891	
O8	−0.4222	−0.2549	
H1	0.0626	0.0121	
H2	0.0626	0.0140	
H(O2′)	0.1350	0.1480	
H3	0.0590	−0.0142	
H4	0.0653	0.0277	
H10	0.0402	0.0007	
H11	0.0513	0.0211	

此外王志中,[99]、刘若庄[100]和他们的合作者用 CNDO 方法计算了 cAMP 和 cGMP。黄天钧和江寿平等[101-105]应用核磁共振技术,对这两分子在溶液中的构象与相互作用进行的研究,表明在一定浓度时它们会自缔合成二聚体,但对它们的二聚体构象并不了解。在上述研究结果基础上又用分子之间相互作用势能方法,对它进行了计算。

(3) cGMP 的双聚体堆积的空间构象[106]

实验证明 cGMP 和 cAMP 在溶液中,它们是会形成二聚体的,但是这二聚体的空间构象是如何安排的,实验未能很好说明。这里以 cGMP 为例在上述分子轨道计算的基础上,应用分子

之间相互作用势能法计算它的二聚体空间构象。这分子是处于顺式构象，又碱基可认为是平面的，其二聚体堆积时有三种可能：面对面，面对背和背对背。通过计算，得到它们的最稳定的堆积模式是 cGMP 分子以顺式构象面对面的堆积模式，两碱基平面之间距离为 0.415 nm，而堆积旋转角为115°。如图9-38和9-39。取向的最优位置主要取决于堆积相互作用能中的静电项贡献。背对背的堆积模式，其关系也是这样。对于 cGMP 双聚体的面对背堆积模式由图可知对于它的极化贡献随旋转角θ的变化是单调的而它的静电项，短程项却随旋转角θ的变化，变化剧烈，这是存在着一点差异的。

总之通过多原子分子间相互作用力理论对于环化鸟苷酸双聚体的各种堆积模式的理论计算及其所得的一系列结果，为我们对于环化鸟苷酸双聚体堆积构象的研究提供了有益的信息，这将有助于对它的结构和功能的研究。

应用分子轨道方法还研究了其它分子[107-109]。国内为了研究药物和生物分子还自编了一整套分子轨道程序[88,89,110-114]。

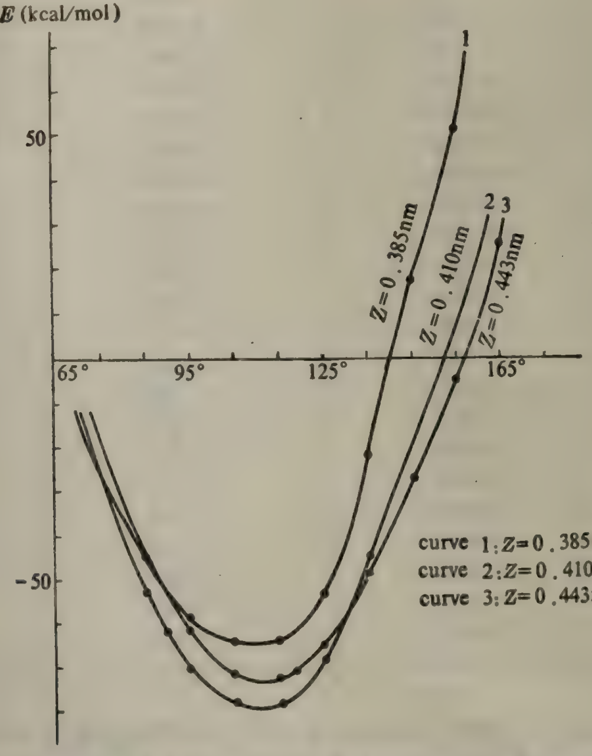


图 9-38 cGMP 分子以顺式构象成“面对面”堆积时，堆积相互作用能与旋转角θ和堆积间距的相互关系

9.3 分子动态学

9.3.1 蛋白质动态学[115]

国内很多人把 Dynamics 和 Kinetics 都无区别地称之为动力学。我们把 Dynamics 称之为

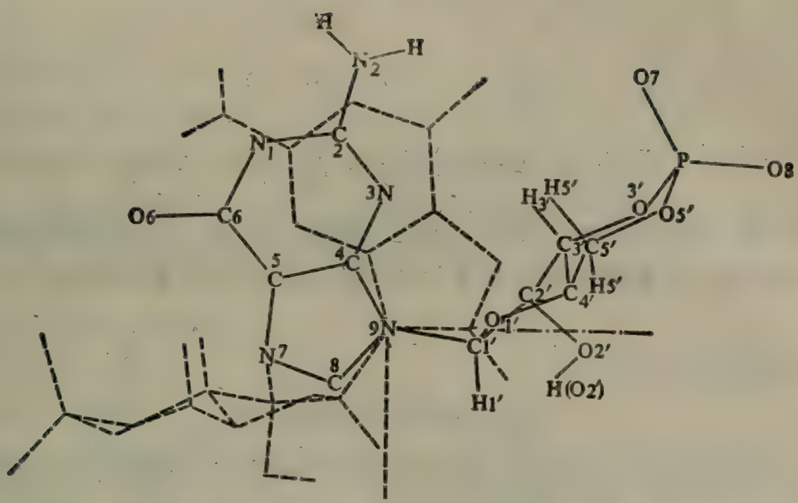


图 9-39 cGMP 双聚体的最稳定构象，堆积模式为“面对面”的堆积，θ为115°，堆积间距Z为 0.415nm。

动态学与 Kinetics 动力学相区别。

生物分子的动态学是近些年来最引人注目的领域之一。经典的观点，从X光得到的信息，认为蛋白质原子是处于固定不动，静止的状态。事实上，蛋白质中原子在常温下是处于不断运动的状态，而X光得到的蛋白质结构，是原子处于平衡位置。生物分子的运动，包括侧链的重新取向和主链的涨落，结构域和亚基的近刚体的运动，到整个多肽链的折叠和伸展，大多数的球蛋白与小分子(配位体、激素、底物和辅酶等)相互作用时，产生构象变化。表9-19列出球蛋白内运动产生的范围。

为了深入了解蛋白质的结构与功能的关系，就必须研究它的内运动。由表9-19可见，蛋白质的运动覆盖着很广的范围，必须用各种方法考察它，而研究蛋白质分子动态学最直接的方法之一是理论计算方法——分子动态学模拟。

表 9-19 球蛋白内运动分类

运动的大小	(300°K)
振幅	0.001 到 10 nm
能量	0.1 到 100 kcal
时间	10^{-14}到 10^{3}sec
运动的形式	
局部的	原子涨落，侧链振荡，环和“臂”位移
大范围	开涨落，折叠和张开
集合的	涨落，孤子(solitoon)和其它非线性运动

虽然量子力学计算能提供小分子的势能面，但分子力学的势能函数是蛋白质分子如此信息的唯一来源。用于蛋白质的能量函数如以前所述，一般是由键、键角、扭转角、Van der Waals相互作用，静电相互作用和氢键，以下式表示：

$$E=\frac{1}{2}\sum_{键}K_b(b-b_0)^2+\frac{1}{2}\sum_{键}K_\theta(\theta-\theta_0)^2$$

$$+\frac{1}{2}\sum_{扭转}K_\phi[1+\cos(n\phi-\delta)]+$$

$$\sum_{非键对}\left(\frac{A}{r^{12}}-\frac{C}{r^6}+\frac{qq_2}{Dr}\right)+\sum_{H键}\left(\frac{A'}{r^{12}}-\frac{C'}{r^{10}}\right) \tag{9.18}$$

这里，(b)键、(θ)键角、(ϕ)二面角、(r)原子间距离，K_b，K_θ，K_ϕ 各是键能参数，A 和 C 是 Lennard-Jones 参数，q_i 是原子电荷，D 是介电常数，A'和 C'是氢键参数，b_0，θ_0，r 和 δ 是几何参考值。

由势能计算所有原子上的力

$$F_i=-\sum_j\nabla E_{ij} \tag{9.19}$$

F_i 是第 i 个原子上的力，E_{ij} 是原子 i 和 j 之间相互作用势。对于质量为 m_i 的原子或分子，按牛顿方程，每一个力产生一个加速度 a_i：

$$F_i=m_ia_i \tag{9.20}$$

分子动态学模拟是人们应用计算机去解蛋白质中原子运动的牛顿方程。现已编成计算机程序，如哈佛大学化学系 Kaplus 等编的 CHARMM，它不仅可进行蛋白质分子的动态学模拟也能进行核酸分子的动态学模拟，分子动态学模拟是到此为止提供最详细和最有趣的蛋白质运动的结果。

Kaplus 应用分子动态学模拟方法模拟了牛胰蛋白酶抑制剂(PTI)，它是有 58 残基，454 个

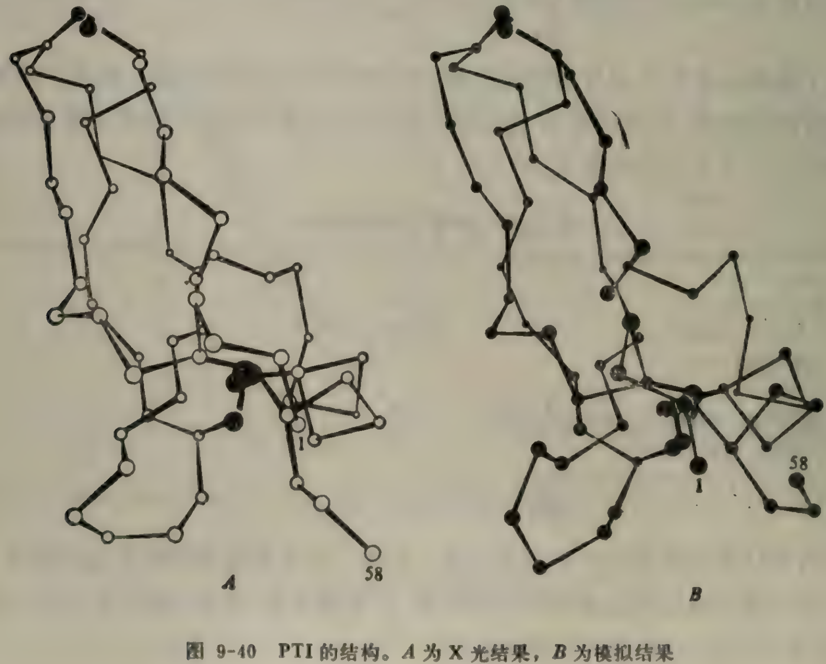


图 9-40 PTI 的结构。A 为 X 光结果，B 为模拟结果

原子的小蛋白质分子，模拟结果表示于图 9-40 中，图中仅表示 α 碳原子和 3 个二硫键，图左边是 X 光衍射的结果，而右边是计算模拟的结果（3 ps 以内的结构），可见这 2 个结构是很类似。

许多的酶和蛋白质分子是由一些多肽链连结的 2 个或更多的结构区域构成。以溶菌酶为例，X 光衍射注意到，当它的活性中心结合抑制剂时，由于围绕着这个活性中心裂缝的 2 个球形的结构域相对位移而有些闭合。Kaplus 计算了溶菌酶裂缝开和闭时构象能的变化，他们的结果表示于图 9-41。图中虚线（带有三角形的）是硬的弯曲势而实线（带圆圈的）是绝热弯曲势。他们还研究了配位体-蛋白质的相互作用，得到满意的结果。



图 9-41　溶菌酶裂缝开（$\theta<0$）和闭（$\theta>0$）的构象变化

9.3.2　酶-底物反应动力学[98]

要了解酶的催化作用，不仅要知道酶分子的结构，也必须研究它与底物分子之间相互作用过程。在溶液中酶和底物分子之间存在着各种相互作用力，但在处理溶液中双分子反应时，没有考虑这些力的作用。随着快速反应动力学的深入研究，出现一个尖锐的问题，即在有些酶与底物的相互作用中，实验所得的二级反应速度常数比按"传统的理论公式"计算结果大一个数量级。为了解决这一问题，我们考虑了分子之间相互作用力：

$$F=-\nabla U \tag{9.21}$$

$$U=U_1+U_2+U_3+U_4+U_5$$

U 为酶-底物分子相互作用势能函数，U_1 是 Van der Waals 势能，U_2 是离子-离子势能，U_3 是离子-偶极势能，U_4 是偶极-偶极势能，U_5 离子-诱导偶极势能。根据牛顿定律，有运动方程。

$$m\frac{\mathrm{d}^2r}{\mathrm{d}t}=-\nabla U-f_1+f_2 \tag{9.22}$$

式中 m 是底物分子的质量，f_1 是底物分子在溶剂中运动所遇到的阻力，f_2 是涨落引起的力。由上述方程可得到酶-底物扩散控制反应新的公式。酶扩散控制反应速度常数最大限度 K_{11m} 和分子力之间关系，由计算机计算结果表示于表 9-20 中。从表可见，考虑了力场后，酶-底物反应速度常数达到了 10^{10}升/摩尔·秒的数量级，从而解释了曾被认为难以解释的实验现象。

表 9-20

U	$U_全$	$U_{1,2,5}$	$U_{1,3,5}$	$U_{1,4}$	$U_{1,5}$	U_1	U_0
K_{11m} 10^{10}升/摩尔·秒	4.056	4.052	3.069	3.058	3.057	3.057	0.20

参考文献

[1] 江寿平：自然杂志，6:912(1982).
[2] 永田亲义著，陶宗晋、江寿平译：《量子生物学入门》，上海科学技术出版社 (1979).
[3] Fraga, S. et al.:"Biomolecular Information Theory", (Elsevier, New York), (1978).
[4] Geisow et al.: "Computing in Biological Sciences", (Elsevier, New York), (1983).
[5] Kotelchuck, D. et al.: Proc. Natl. Acad. Sci. U. S. A., 62:14(1969).
[6] Robson, B. et al.: J. Mol. Biol., 58:237(1971).
[7] Lim, V. I.: J. Mol. Biol., 88:857 (1974).
[8] Chou, P. Y. et al.: "Adxances in Ezymology", Vol. 47, (John Wiley, New York), (1978).
[9] Schulz et al.: Nature, 250:140(1974).
[10] 刘志平、江寿平：生物化学与生物物理学报，17:657(1985).
[11] 刘志平、江寿平：科学通报，30:625(1985).
[12] Hilda, C. et al.: FEBS Letters, 150:247(1982).
[13] 甘人宝、李载平：中国科学，B辑，(12):1114(1983).
[14] 王亚辉：《分子免疫》，科学出版社，(1985).
[15] Tomas, P. H. et al.: Proc. Natl. Acad. Sci. U.S.A., 78:3824(1981).
[16] Hayes, T. G.: Biochem. Biophys. Res. Commun., 95:872(1980).
[17] Zoom, K.C. et al.: "Handbook of Experimental Pharmacology", Vol. 71, (Springer-Verlag, Berlin), (1984).
[18] 程玉华等：生物化学与生物物理学报，17:8(1985).
[19] 王德宝等：《核酸》，科学出版社，(1986).
[20] Roberts, R. J.: Nucleic Acids Res., 12:1 (1984).
[21] Roberts, R. J.: Nucleic Acids Res., 14:1 (1986).
[22] 乐树云，江寿平：科学通报，28:1521(1983).
[23] 乐树云，江寿平：生物化学与生物物理学报，16:161(1984).
[24] 乐树云，江寿平：生物化学与生物物理学报，16:365(1984).
[25] 乐树云，江寿平：生物化学与生物物理学报，16:425(1984).
[26] 乐树云，江寿平：生物化学与生物物理进展：(6):64(1982).
[27] 乐树云，江寿平："The China-Japan Bilateral Symposium on Biophysics", 1985.
[28] 乐树云，江寿平：生物科学动态，(3):1(1987).
[29] 钱雯、乐树云：生物化学杂志，2:75(1986).
[30] 黄玉明：生物化学与生物物理进展，(1):2(1987).
[31] Tinoco, I. Jr. et al.: Nature, 230:326(1971).
[32] Nussinov, R. et al.: Proc. Natl. Acad. Sci. U.S.A.,77:630(1980).
[33] Zuker, M. et al.: Nucleic Acids Res., 9:133(1981).
[34] 乐树云、江寿平：科学通报，29:302(1984).

[35] 乐树云、江寿平：生物物理学报 2:76(1986).
[36] Noller, N. F. et al.: Science, 212:403(1981).
[37] Noller, H. F.: Ann. Rev. Biochem., 53:119(1984).
[38] 乐树云、江寿平：生物化学与生物物理学报，17:557(1985).
[39] 乐树云、江寿平：生物化学与生物物理学报，17:664(1985).
[40] 乐树云、江寿平：生物化学与生物物理学报，17:722(1985).
[41] 乐树云、江寿平：生物物理学报，4:227(1987).
[42] Nazar, R. N. et al.: J. Biol. Chem., 250:8591(1975).
[43] Ford, R. J. et al.: Eur. J. Biochem., 37:199(1978).
[44] Luoma, G. A. et al.: Proc. Natl. Acad. Sci. U.S.A., 75:4901(1978).
[45] 刘次全等：动物学研究，5:79(1984).
[46] Siebenlist, U. et al.: Cell, 20:269(1980).
[47] Ramachandran, G. N. et al.: Adv. Protein Chem., 23:283(1968).
[48] Scheraga, H. A.: Adv. Phys. Organ. Chem., 6:103(1968).
[49] Momany, F. A. et al.: J. Phys. Chem., 79:2361(1975).
[50] Pullman, B. et al.: Adv. Protein Chem., 28:348(1974).
[51] Hillier, I. H.: J. Theor. Biol., 76:83(1979).
[52] Sasisekharan, V.: Jerus. Symp. Quant. Chem. Biochem., 5:247 (1974).
[53] Govil, G. et al.: J. Theor. Biol., 33:399(1971).
[54] Saran, A. et al.: Theor. Chim. Acta, 30:31(1973).
[55] 吴伟雄、江寿平等：生物化学与生物物理学报，13:495(1981).
[56] 乐树云、江寿平等：生物化学与生物物理学报，13:433(1981).
[57] 乐树云、江寿平：结构化学，2:57(1983).
[58] 江寿平、乐树云等：生物化学与生物物理学报，13:485(1981).
[59] 乐树云、江寿平：分子科学学报，2:87(1982).
[60] 郝鸣鸿、杨频等：生物化学与生物物理学报，18:217(1986).
[61] 江寿平、乐树云：分子科学学报，2:35(1982).
[62] 乐树云、江寿平：科学通报，27:172(1982).
[63] 倪向善、陈润生：中国科学，B辑，(1):45(1985).
[64] Alagona, G. et al.: Int. J. Pept. Prot. Res., 5:251(1973).
[65] Del Bene, J. E.: J. Chem. Phys., 62:1961(1975).
[66] Hylton, J.: J. Am. Chem. Soc., 98:7198(1976).
[67] Vo Velle, F. et al.: Inter. J. Pept. Prot. Res., 13:435(1979).
[68] Pullman, A. et al.: Quant. Rev. Biophys., 7:530(1974).
[69] Pullman, A. et al.: Theor. Chim. Acta, 50:317(1978).
[70] Nemethy, G. N. et al.: Quant. Rev. Biophys., 10:239(1978).
[71] Ghelis, C. et al.: "Protein Folding", (Acad. Press, New York), (1982).
[72] Levitt, M.: J. Mol. Biol., 104:59(1976).
[73] Scheraga, H. A. et al.: Quant. Chem. Program Exchange, 11:286(1975).
[74] Nemethy, G. N. et al.: J. Phys. Chem., 87:1883(1983).
[75] Warshel, A, et al.: Quant. Cham. Program Exchange, 11:247(1974).
[76] Fraga, S.: Prediction of Protein Tertiary Structure, University of Alberty, (1985).
[77] 刘志平、江寿平：科学通报，31:62(1986).
[78] 刘志平、江寿平："The China-Japan Bilateral Symposium on Biophysics", (1985).
[79] 刘志平、江寿平等: Biopolymers, 25:s169(1986).
[80] 钮经义等：科学通报，27:1004(1982).
[81] 陈凯先等：分子科学学报，2:103(1982).
[82] 王志中等：《半经验分子轨道理论与实践》，科学出版社，(1981).
[83] Dion, D. R.: QCPE, 368(1978).
[84] Boyd, D. B. et al.: J. Theor. Biol., 25:403(1969).

[85] Alving, R. E. et al.: ibid, 34:199(1972).
[86] 张学忠等：分子科学学报，2:75(1982).
[87] 江寿平、吴伟雄：科学通报，27:1071(1982).
[88] 吴伟雄、江寿平：生物化学与生物物理进展，(2):73(1982).
[89] 宣建成、江寿平：中国科学，B辑，(7):681(1986).
[90] Drummond, M. L. J.: Progress in Biophysics & Molecular Biology, 47:1(1986).
[91] Hinze, J.: Chem. Phys., 26:213(1974).
[92] 宣建成、江寿平：分子科学学报，5:63(1987).
[93] 江寿平、宣建成：中国科学，B 辑，(6):609(1987).
[94] 江寿平、宣建成：科学通报，30:1819(1985).
[95] 江寿平、宣建成：科学通报，32:182(1987).
[96] 江寿平、吴伟雄：分子科学学报，2:49(1982).
[97] 江寿平等：生物物理学报，3:65(1987).
[98] 江寿平：中国科学，(4):329(1977).
[99] 王志中等：高等学校化学学报，5:432(1984).
[100] 刘若庄等：生物化学与生物物理学报，18:474(1986).
[101] 黄天钧、江寿平等：生物化学与生物物理学报，14:225(1982).
[102] 黄天钧、江寿平等：生物化学与生物物理学报，15:297(1983).
[103] 黄天钧、江寿平等：分子科学学报，4:107(1984).
[104] 黄天钧、江寿平：结构化学,2:175(1983).
[105] 江寿平等：XII Inter. Confer. on Magnetic Resonance in Biological Systems, (1986).
[106] 江寿平等：结构化学，1:83(1982).
[107] 吴伟雄、江寿平：分子科学与化学研究，3:143(1983).
[108] 汪信、忻新泉：分子科学与化学研究，2:15(1982).
[109] 吴伟雄、忻新泉：结构化学，1:85(1982).
[110] 刘次全等，科学通报
[111] 陈凯先等：分子科学学报，1:65(1981).
[112] 吴吉安等：化学通报，(11):13(1984).
[113] 吴伟雄、忻新泉等：南京大学学报(化学专刊)，139(1985).
[114] 吴伟雄、江寿平：自然杂志，4:636(1981).
[115] Karplus, M. et al.: Ann. Rev. Biochem., 52:263(1983).

中国科学院上海生物化学研究所生物大分子构象组
陈　洪

第10章　金属离子生物学功能的量子生物学研究

10.1　概　　述

在生物学基础课程中，有关金属离子的生物学功能已经作了初步的介绍。我们知道，在生物体内，金属离子参与了重要的生命过程，如

$$Fe^{3+} + e \longrightarrow Fe^{2+}$$

$$Fe^{2+} - e \longrightarrow Fe^{3+}$$

的反应正是生物体的重要生命过程之一[1]。最近十几年发展起来的"生物无机化学"(或"无机生物化学")已将金属离子在生物系统中的作用列为主要研究对象。但从发展趋势看，多学科的协同研究将会更加明显。在这一章里，我们将着重讨论从量子生物学角度研究金属离子生物学功能的问题。

有人认为，生物无机化学的理论前景的基础是，应用现代量子化学概念来解释生物体及其模型[2]。事实上，应用量子生物学的理论和方法研究金属离子与生物分子(包括生物大分子)间的相互作用，已经作了一些有意义的工作[3-12]。然而，由于生物分子大而复杂，它和金属离子的作用又是在水或油脂介质中进行的，这些作用不仅涉及到一些基团电子运动的变化，而且涉及到多种化学键的集体效应，特别是还涉及到生物大分子的构象变化等等。因此，要真正严格计算金属离子与生物分子间的相互作用，生物金属配合物的电子结构(见后)等等，势必还涉及到大量原子的动态过程。尽管严格的计算尚难实现，但由于计算方法的不断改进和完善，新的计算方法的出现，以及高速大容量电子计算机的发展，尤其是理论计算同实验研究的紧密配合，使得有关"金属离子在生物体中作用"的量子生物学研究不断取得进展。

10.2　金属离子生物学功能的配位场理论[13—15]

大量的研究已经证实，多数与金属离子有关的生物学过程实际上是通过金属离子与生物分子的相互作用而发生的，这些金属离子通过引起构象的改变或在生物大分子中产生活性中心等来稳定、分解、开关、调节和控制这些生物大分子[16]。但是，在生物体内，金属离子的作用往往又是通过与生物活性配体(见后)形成配位化合物(过去我们称之为络合物，以下简称配合物) 来体现的。然而，在解释配合物的结构和性质上，一般认为比较成功的理论是所谓"配位场理论"。该理论源于 Bethe 和 Van Vleck 等的工作。

配位场理论是在晶体场理论的基础上和分子轨道理论相结合而发展起来的一种理论，因而亦有二者的优点。以下我们将分作几个问题来讨论。

10.2.1　与金属离子相互作用的生物配位体*

在生物体内可与金属离子配位的配体，大体上可以分为三类：

第一类为复杂的生物大分子，如蛋白质、核酸、多糖、磷脂以及它们的各级降解产物（如氨基酸、肽、碱基、核苷、核苷酸和低聚糖等）。

第二类为简单分子，如水、氧、胺类、羧酸、某些碳水化合物及其有关衍生物。

第三类是一些简单的有机、无机离子，如有机酸根、Cl^-、SO_4^{2-}、HCO_3^-、CO_3^{2-}、$H_2PO_4^-$等等。

除此而外，其它生物配位体还可举出维生素和激素等等。

在上述配位体中，大分子配位体与小分子配位体相比，在与金属形成配合物方面有以下特点：

(1) 大分子配位体是一个多配位部位的配位体。例如蛋白质分子中的每个氨基、游离羧基和个别氨基酸残基侧链的巯基、杂环氮等都可参与对金属离子的竞争。以金属离子为配位中心的活性部位就是由这种竞争产生的。

(2) 大分子配位体具有高级结构。它的折叠卷曲在组成金属离子的配位环境方面和形成反应中心附近的微环境方面起着决定性作用。

(3) 由于大分子配位体的刚性与柔性相结合，动态与静态相结合，从而使它与金属离子起作用时，整个分子发生变动。

(4) 大分子配位体具有特殊的溶解性，能够携带较多的金属离子而保持"溶解"状态。

在生物体内所作的研究表明，有时金属离子本身并没有某种生物活性，或者活性不够，而只有当其与具有特定结构的配位体作用形成配合物之后，才表现出特定的活性。例如 Fe^{2+} 本身不能氧合，只有当它与原卟啉 IX 和特定结构的蛋白质链结合成血红蛋白后才能氧合。

需要指出，我们在这一章里要探讨的主要问题是，金属离子与生物配位体间的相互作用，因为没有这种相互作用，就不可能在生物体内保持分子的天然结构及其正常功能。在此基础上，进而讨论一些生物金属配合物的生物学功能。

10.2.2　配合物的晶体场理论

晶体场理论实质上是一种静电理论，它把配合物中中心离子与配位体间的相互作用，看作是类似于离子晶体中正负离子间的相互作用，所以叫晶体场理论。

晶体场理论的基本要点是，它把配位场对中心金属离子的作用看成是一种静电性质的微扰。微扰的一个最重要的结果是原来自由金属离子的那些简并状态分裂为两个或两个以上的非等价状态——简并能级的分裂。作为一级近似，我们一般只考虑在配位体场的影响下，金属离子 d 轨道能级的分裂。

实验研究也证实，生物金属配合物的许多性质，都与配合物中金属离子 d 能级的分裂有关。

我们知道，中心金属离子的 d 轨道有五种，即 d_{xy}、d_{yz}、d_{xz}、$d_{x^2-y^2}$和 d_{z^2}。这五种 d 轨道的能

* 所谓生物配位体是指生物体内能同金属离子相互作用形成配合物的离子和分子。

量本来是相等的，但当生物配位体接近金属离子时，由于配位体静电场的作用，使五种 d 轨道的能量发生不同程度的变化。例如有六个配位体，分别沿 $\pm x, \pm y, \pm z$ 的方向向中心离子接近而形成正八面体配合物（如图 10-1 所示的铁卟啉配合物模型）时，d_{z^2} 和 $d_{x^2-y^2}$ 轨道就与配位体处于迎头相碰的状态，在这些轨道上的电子，由于受配位体的静电排斥作用，因而能量较高，而 d_{xy}、d_{xz}、d_{yz} 轨道的电子云正好插在配位体的空隙中间（如图 10-2），因而这三个轨道的能量较低。于是，本来能量相等的 d 轨道便分裂为二组：一组是两个能量较高的 d_{z^2} 和 $d_{x^2-y^2}$ 轨道，习惯上叫做 e_g 轨道，另一组是三个能量较低的 d_{xy}、d_{xz} 和 d_{yz} 轨道，习惯上叫做 t_{2g} 轨道。这两组的能级差 Δ_0 叫做分裂能。如图 10-3 所示，能级差 Δ_0 既与中心离子有关，也与配位体有关。需要注意的是，

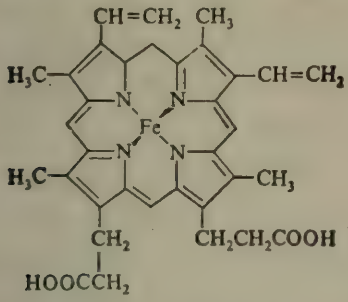


图 10-1 一种铁卟啉配合物

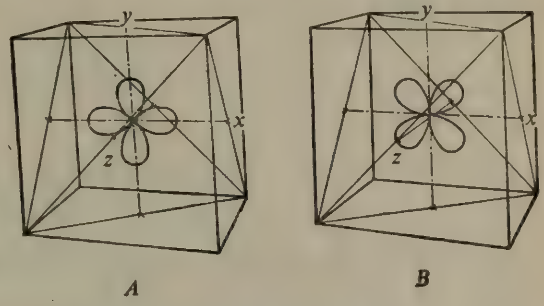


图 10-2 在八面体场中（A）e 轨道 $d_{x^2-y^2}$ 和（B）t 轨道 d_{xy} 的取向

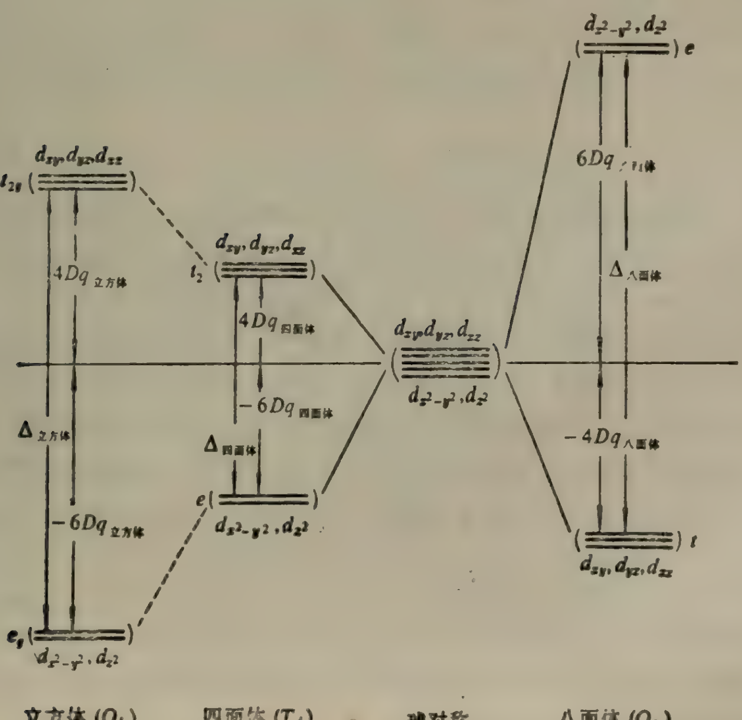


图 10-3 在八面体配位场中 d 轨道的分裂

在配位体场的影响下，五个 d 轨道的能量都有提高，只不过 e_g组轨道提高更多。

设八面体配合物的分裂能 $\Delta_{八面体}$为

$$\Delta_{八面体}=E(e_g)-E(t_{2g})=10\ Dq_{八面体} \tag{10.1}$$

从对称性考虑，t_{2g}轨道是三重简并，e_g是二重简并，并取全部充满时总能量为零．则有

$$4\ E(e_g)+6\ E(t_{2g})=0 \tag{10.2}$$

故 $E(t_{2g})=-4\ Dq_{八面体}$．$E(e_g)=6\ Dq_{八面体}$

如果认为配位体和中心离子之间是静电相互作用，则由量子力学微扰理论计算的近似结果为

对点电荷模型：

$$10\ Dq_{八面体}=\frac{5\ eq\bar{a}^4}{3\ \bar{r}^5} \tag{10.3}$$

对偶极子模型：

$$10\ Dq_{八面体}=\frac{5\ e\mu\bar{a}^4}{\bar{r}^6} \tag{10.4}$$

式中，e 为电子电荷。由(10.3)和(10.4)式可见，能级分裂Δ_0的大小决定于：(1) 配位体场的强弱(反映在配位体电荷 q 或偶极矩 μ)，(2) 离子和配位体中心间的间距 $\bar{r}$，(3) 金属离子的本性(反映在 d 电子轨道半径四次方的平均值 $\bar{a}^4$)。

对其它构型也可作类似讨论，假定中心离子和配位体的距离保持不变，则有关系

$$\Delta_{四面体}=-\frac{4}{9}\Delta_{八面体}$$

$$\Delta_{立方体}=2\ \Delta_{四面体}=-\frac{8}{9}\Delta_{八面体} \tag{10.5}$$

分裂能还给配合物带来额外的稳定能，我们称之为晶体场稳定化能(CFSE)。它是将 d 电子从未分裂的 d 轨道 E_s能级进入分裂的 d 轨道所产生的总能量下降值(参见图 10-3)。下降愈多，即 CFSE 越大，配合物也就越稳定，所以配合物的稳定性可用 CFSE 的大小来衡量。

总之，晶体场理论从一个静电的晶体场模型，提出了 d 轨道的能级分裂和稳定化能等概念，用以说明了许多化学事实，例如配合物的立体结构、磁性、吸收光谱和配合物在溶液中的稳定性等。所以这一理论从五十年代初以来有较大的发展。但是同时它又有严重的缺陷，模型也过于简单．仅着眼于金属离子和配位体之间的静电作用，而未考虑它们之间的轨道重叠。其实，研究已经表明，当金属离子与生物配位体相互作用时，轨道的重叠是不可忽视的。因此，对于这个理论的修正，就是要考虑到金属离子的轨道与配位体的轨道重叠。

除用晶体场理论解释配合物中心金属离子 d 轨道的能级分裂，用其它方法，如群论方法也可对 d 轨道能级分裂的情况作出解释。但是，所需的群论知识较多。然而，由王荣顺等建立的指在论述 d 轨道分裂的对称性方法，仅用对称性知识，对外场作用下 d 轨道能级分裂的情况，作出既严格又可被读者接受的讨论(参见文献[89])。

10.2.3 配合物的分子轨道理论

在上面讨论的晶体场理论能简单明了地解释不少实验事实，其作用是应该肯定的，正因为如此，所以该理论至今仍在运用。但是，人们不禁要问，如果处理非极性配位体（譬如非极性的生物配体），那又该如何解释呢？显然，在这种情况下，配位体与中心金属离子间的轨道重叠作用就成了配合物成键的一个重要原因。一些实验，如顺磁共振实验结果也证明了 d 电子会离域到配位体上去，配合物的光谱（特别是含有 π 键或 d 轨道的配位体）也说明要考虑共价特性所引起的变化。鉴于这些原因而发展的配合物分子轨道理论，在实际上已经证明可以说明这方面的问题。

我们知道，共价键依据其对称性质可分为 σ 键和 π 键。前者绕键轴转动 键波函数不改变符号，后者绕键轴转动 180°键波函数改变符号。同样地，配价键中也有 σ 配键和 π 配键。

根据分子轨道理论，由中心离子M和配位体 L 的价电子轨道互相重叠，可得配合物的分子轨道。现以八面体配合物 ML_6为例简述之。

设中心金属离子M处在直角坐标系 X、Y、Z 的原点，六个配位体位于坐标轴上。金属离子有九个原子轨道可以形成分子轨道，它们是：$ns, np, (n-1)d$。这九个轨道可分为两组，一组是可与配位体生成以轴为对称的 σ 分子轨道。另一组是可与配位体生成面对称的 π 分子轨道。根据坐标系的选择，这两组轨道分别为：

σ：$s, p_x, p_y, p_z, d_{z^2}, d_{x^2-y^2}$

π：d_{xy}, d_{yz}, d_{xz}

由原子轨道线性组合成分子轨道，必须满足对称匹配的条件。为此，首先把这六个配位体的 σ 轨道线性组合，以构成配位体的群轨道，然后再把它与金属离子的相应对称轨道线性组合成 σ 分子轨道。

中心金属离子的六个原子轨道与配位体的六个群轨道线性组合形成 12 个分子轨道，其中 6 个是成键轨道，6 个是反键轨道。按照分子轨道理论，八面体配合物的分子轨道应由 15 个轨道线性组合成 15 个分子轨道（其中，d_{xy}, d_{yz}和 d_{xz}不能形成 σ 键，而能形成 π 键），其分子轨道图如图 10-4 所示。其中 a 表示没有简并的状态，e 表示二重简并，t_1和 t_2分别表示两种不同状态的三重简并态，g 表示其波函数是中心对称的，而 u 表示是中心反对称的。分子轨道各能级的相对高低是由量子化学计算和光谱实验确定的。

由分子轨道理论知，当一个分子轨道的能级，接近于用来建立它的原来两方轨道中一方的能级时，这个分子轨道的性质就和接近它的那个轨道相似。从图 10-4 还可以看出，配位体的 σ 轨道能级比中心离子的 d 轨道能级低，生成的六个成键轨道比配位体孤对电子的 σ 轨道能量低，因而六个配位体的 12 个孤对电子将进入这些成键轨道中，相当于生成共价配键。

正如前面所述，在一些配合物中，金属离子与配体间除有 σ 键合外，还可以有 π 键合。对于八面体配合物，在金属离子与配位体间的 σ 键合中，金属离子的 t_{2g} 轨道为非键轨道。但如果有 π 键，那就要考虑配位体与金属离子的 t_{2g} 轨道重叠成 π 键合。图 10-5 表示金属离子的t_{2g}轨道与配位体 p 轨道重叠的情形。

以上我们简述了八面体配合物的情形，对于其它配合物的分子轨道就不再一一叙述，而在后

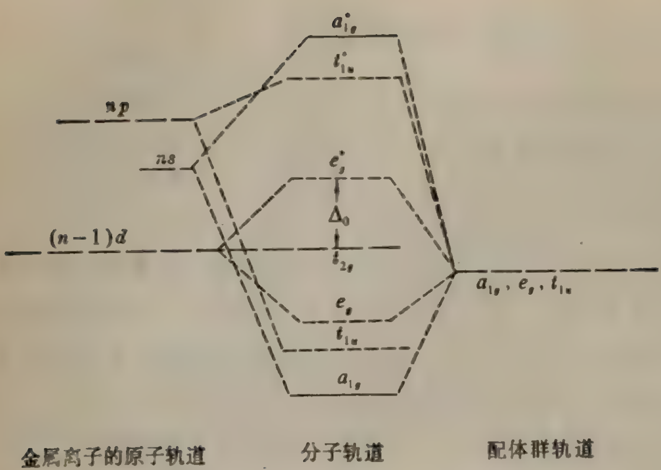


图 10-4 八面体配合物中 σ 分子轨道的能级

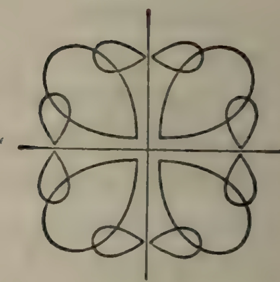


图 10-5 八面体配合物中金属离子 t_{2g}轨道与配位体 p 轨道的重叠

续部分我们将详尽地讨论一些生物金属配合物的分子轨道计算，以及由此而得到的一些结果。

10.2.4 配位场理论的应用

研究实践表明，配位场理论可以用来解释配合物的很多物理和化学性质。

运用配位场理论已经搞清楚了 d 过渡金属配合物分为高自旋与低自旋的原因。当一对电子处于同一轨道中时，由于两个电子处于同一轨道比在两个不同轨道中的静电斥力大，因而它们之间将产生很强的排斥作用，这种作用称为电子成对能。如果轨道的能级分裂值（Δ_0 或 D_q 值）大于成对能，则电子就填充到能量最低的轨道（在八面体配合物中能量最低轨道即是 t_{2g} 轨道）。这样，就形成低自旋配合物。如果轨道的能级分裂值不太大，则电子将占有两个能级，即尽量使不成对的电子数为最大，从而形成高自旋配合物（见表 10-1）。

表 10-1 3d-生物金属的自旋状态特征

金属离子	电子构型	电子在t_{2g}和e_g轨道中的分布及不成对电子数	
		高自旋络合物	低自旋络合物
Mn^{2+}	d^5	$(t_{2g})^3\ (e_g)^2$5	$(t_{2g})^5$1
Fe^{3+}	d^5	$(t_{2g})^3\ (e_g)^2$5	$(t_{2g})^5$1
Fe^{2+}	d^6	$(t_{2g})^4\ (e_g)^2$4	$(t_{2g})^6$0
Co^{3+}	d^6	$(t_{2g})^4\ (e_g)^2$4	$(t_{2g})^6$0
Co^{2+}	d^7	$(t_{2g})^5\ (e_g)^2$3	$(t_{2g})^6\ (e_g)^1$1

运用配位场理论可以解释配合物的电子光谱，说明配合物的磁性以及用来对配合物反应机理进行研究。

从配位场理论可以推测，金属将优先和分裂能值（或 D_q 值）大的配位体作用，因为这样具有

较大的晶体场稳定能。一般氨及胺的 D_q 值比水或其它成键原子是氧的配位体要大。因此，过渡金属将优先地和氮而不是氧成键。对于非过渡金属、稀土金属和半充满的 Fe^{3+} 及 Mn^{2+} 等离子，将倾向于优先和氧而不是和氮成键。这些理论研究结果已经得到实验的证实。

我国唐敖庆等[17] 对配位场理论作了许多有意义的研究工作。他们提出了改进的弱场方案、使计算作用能矩阵化的方法标准化，并应用改进的弱场方案计算了 d^4 及 d^6 组态正八面体配合物的能量矩阵元。这些工作在国内外都占有相当的地位。

总之，从发展趋势看，配位场理论已经成为解释金属配合物结构和性质方面很有威力的理论。

10.3 金属离子与核酸间的相互作用

过去曾经认为，核酸与金属阳离子的结合，系发生在核酸的磷酸基团的负电荷部位。后来的研究证实，金属阳离子在较大程度上是结合在含氮碱基上，即与腺嘌呤和鸟嘌呤作用生成配位化合物，因为这两种碱基分子为金属螯合作用提供了合适的空间构型。图 10-6 表示一种稳定构型。

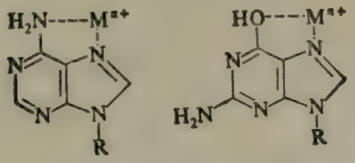


图 10-6 金属离子与腺嘌呤和鸟嘌呤环螯合的可能结构

从这种稳定构型看，金属离子跟腺嘌呤和鸟嘌呤碱基形成配合物包括两种碱基的 N_7，以及跟 C_6相联结的位于环外的杂原子，也许这是形成有五个成员的一种特别稳定的螯合环的最适位置。

10.3.1 金属离子与核酸碱基相互作用的最优配位模式

近几年，江寿平等[5]根据分子中原子和离子之间非键相互作用的半经验势函数，并应用了 BFGS 变尺度法进行最优化处理，研究了 Na^+与 DNA 中 G、C、A、T、G-C、A-T碱基对和双联体 $\downarrow{}^{AT}_{GC}\uparrow$、$\downarrow{}^{AT}_{CG}\uparrow$ 等的相互作用，得到了它们各自的最优配位模式。Na^+与胞嘧啶、鸟嘌呤、腺嘌呤、胸腺嘧啶、G-C和 A-T 碱基对相互作用的最优配位模式如图 10-7、10-8、10-9、10-10、10-11 和 10-12 所示。

表 10-2 列出了江寿平等运用离子-分子相互作用势能法计算的 Na^+与碱基 A、C 的相互作用能，并同 ab initio SCF 法计算结果[18]作了比较。

比较表明，两种方法计算结果在数值上是接近的，相对误差约在 8.1～13.3%，与此同时，我们还可从图 10-13 与图 10-7 至 10-12 的比较中看到这两种方法的接近程度。此外，由 Coulter

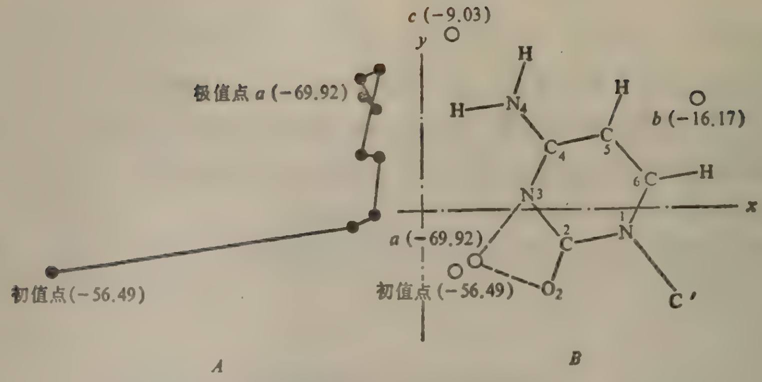


图 10-7 A. 由初值点到极值点 a 的Na^+最佳游动路径，比例 1:20
B. Na^+与胞嘧啶相互作用的最优位置和相互作用能(kcal/mol)

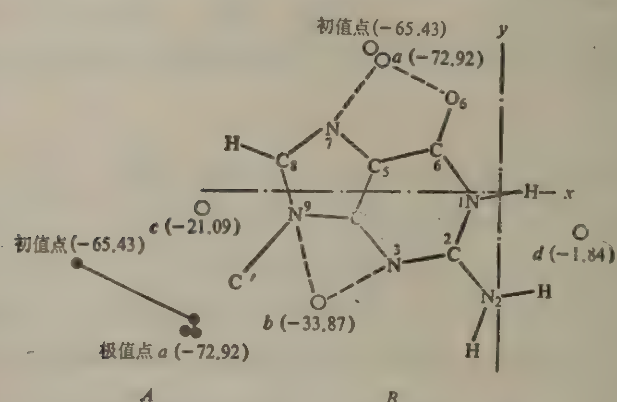


图 10-8 A. 由初值点到极值点 a的最佳游动路径，比例 1:10
B. Na^+与鸟嘌呤相互作用的最优位置和相互作用能(kcal/mol)

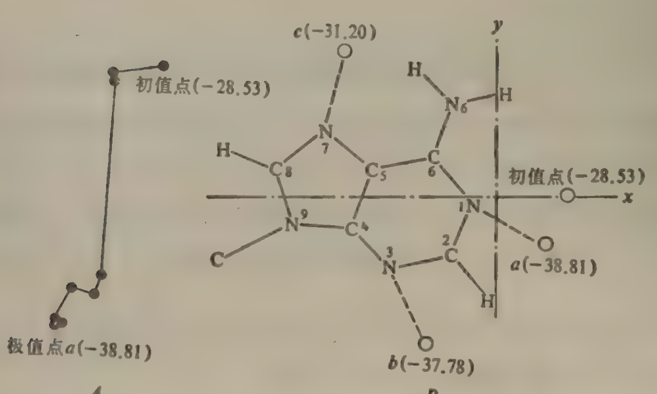


图 10-9 A. 由初值点到极值点 a 的 Na^+ 最佳游动路径，比例 1:5
B. Na^+与腺嘌呤相互作用的最优位置和相互作用能(kcal/mol)

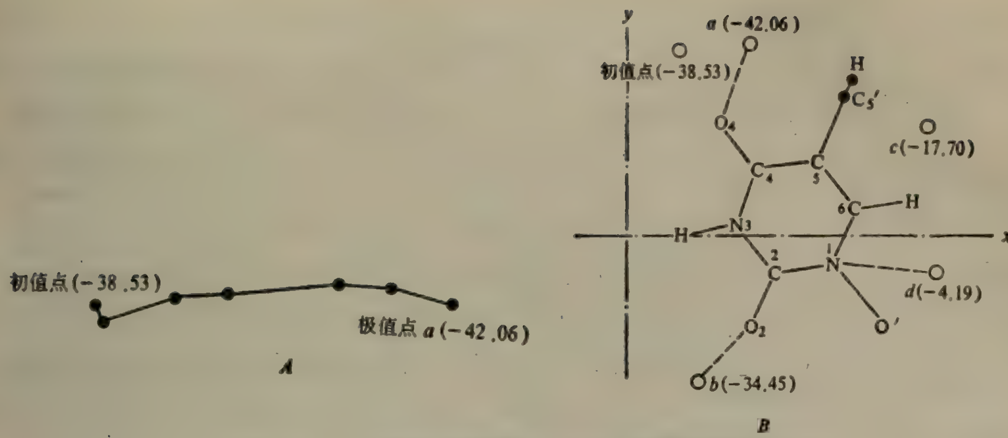


图 10-10 *A*. 由初值点到极值点 *a* 的 Na^+ 最佳游动路径，比例 1:5
B. Na^+ 与胸腺嘧啶相互作用的最优位置和相互作用能(kcal/mol)

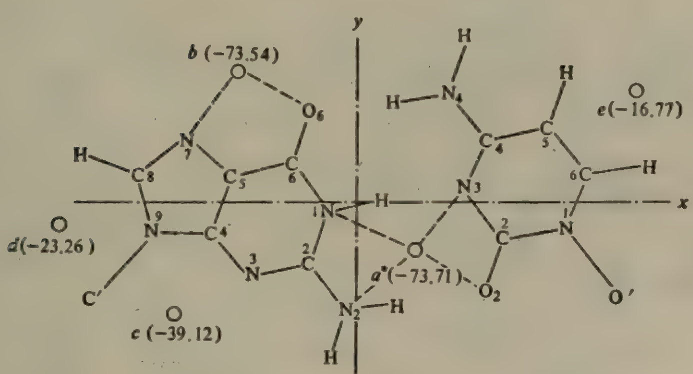


图 10-11 Na^+ 与碱基对 G-C 相互作用的最优位置和相互作用能(kcal/mol)
* 极值点 *a* 的 *z* 坐标为 $z=\pm 0.9879$

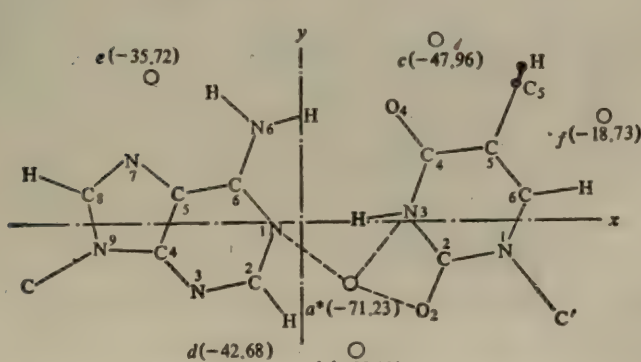


图 10-12 Na^+ 与碱基对 A-T 相互作用的最优位置和相互作用能(kcal/mol)
* 极值点 *a* 的 *z* 坐标为 $z=\pm 0.5914$

表 10-2　离子-分子相互作用势能法与ab initio SCF法计算N_a^+与碱基C、A相互作用能结果的比较

	Na⁺与碱基分子的配位位置(Å)	相互作用能(kcal/mol)		相对误差
		ab initio	江的工作	
Na⁺与胞嘧啶相互作用	$d_{Na^+}\cdots N(3)=2.30$, $d_{Na^+}\cdots O(2)=2.15$	−51.70	−56.49	9.26%
Na⁺与腺嘌呤相互作用	$d_{Na^+}\cdots N(1)=2.15$, $\angle Na^+N(1)C(6)=125°$	−26.40	−28.53	8.07%
	$d_{Na^+}\cdots N(3)=2.15$, $\angle C_2N_3Na^+=\angle C_4N_3Na^+$	−24.00	−26.06	8.58%
	$d_{Na^+}\cdots N(7)=2.15$, $\angle Na^+N_7C_8=135°$	−21.30	−24.14	13.33%

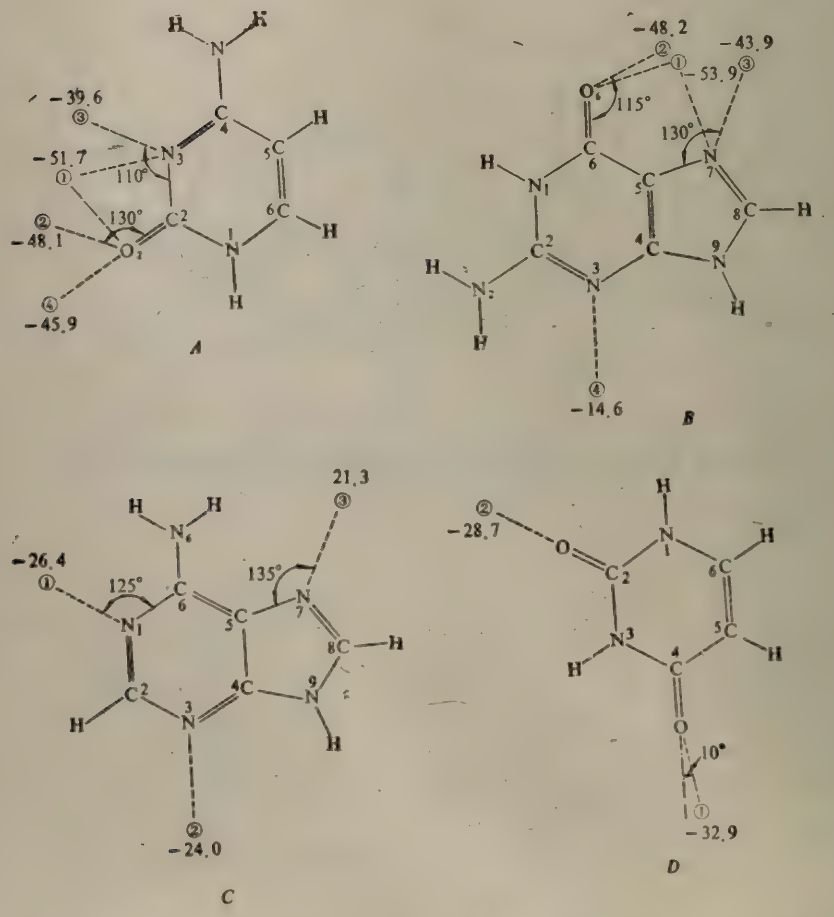


图 10-13　ab initio SCF 法计算得到的 Na⁺ 与碱基相互作用的最优位置和相互作用能(kcal/mol), *A*. 胞嘧啶, *B*. 鸟嘌呤, *C*. 腺嘌呤, *D*. 尿嘧啶[19]

等[20]对于 Na(β)-2′-3′-环化胞苷酸等的 X 射线晶体结构测定结果也表明，Na^+与 G、C 的配位模式与江寿平等由计算得到的最优配位模式是一致的。

根据图 10-7 至 10-13，我们可以推测出，当Na^+与核酸碱基、碱基对和双联体 $\downarrow {}^{AT}_{GC} \uparrow$、$\downarrow {}^{AT}_{CG} \uparrow$ 相互作用形成配位化合物时，Na^+的可能最优配位位置。

唐明生和刘若庄等[21]用 CNDO/2 方法对环腺苷酸(cAMP)与 Cu^{2+} 的配位作用也进行了理论研究。我们知道，cAMP 是存在于生物体内的一种生物小分子，它作为第二信使在生物体内起着广泛的调节作用。由于它在活细胞中起着重要的调控作用，因而日益受到人们的注视。近年来，有关 cAMP 的理论和实验研究更加广泛、深入，探讨其与金属离子配位作用的理论计算也相继展开。从唐明生等的计算结果看，当 cAMP 分子与金属离子相互作用时，由于其碱基上的 N_7 原子具有较大的络合能力(从图 10-14 可以看出，N7、N1、N3原子都具有负的净电荷。N7原子为－0.2160；N1原子为－0.1110；N3原子为－0.1296，其中 N7原子具有的负电荷最多)，因而反应也将优先在 N7原子上进行。

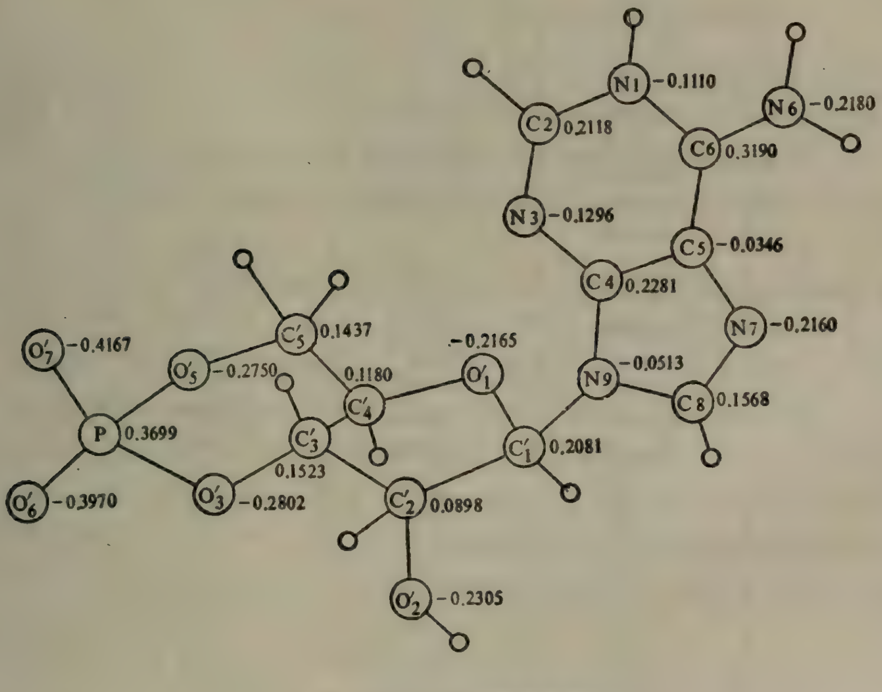


图 10-14 cAMP 的净电荷分布图

进一步比较 cAMP 分子与 cAMP 的Cu^{2+}配合物的电荷分布(表 10-3) 可以看出，cAMP分子中P原子的净电荷为＋0.3699，O6′，O7′，O5′，O3′的净电荷分别为－0.3970，－0.4167，－0.2750，－0.2802。而 cAMP 的 Cu^{2+}配合物中的P原子的净电荷为＋0.5990，O6′，O7′，O5′，O3′ 的 净 电 荷分别为－0.1159，＋0.0016，－0.2378，－0.2470。由此可见，金属离子 Cu^{2+}与cAMP分子的配位相互作用将使得P原子上的净正电荷增加0.2991，且使得 P原子周围的氧原子上所带的负电荷大大减少。这必然导致阴离子对 P 原子的亲核进攻变得更加有利。因此，根据

表 10-3 净电荷分布

净电荷 \ 分子或分子片断 / 原子	cAMP	cAMP与Cu^{++}的络合物	碱基片断	磷酸戊糖环片断
C4′	0.1180	0.1161		0.1090
O1′	−0.2615	−0.2288		−0.2247
C1′	0.2081	0.0079		0.1269
C2′	0.0898	0.0984		0.1130
O2′	−0.2305	−0.2100		−0.2490
H(O2′)	0.1189	0.1515		0.1392
H(C2′)	−0.0036	0.0111		0.0090
C3′	0.1523	0.1570		0.1579
H(C3′)	−0.0266	−0.0211		−0.0058
H(C1′)	−0.0004	0.0345		0.0151
H(C4′)	0.0173	0.0411		0.0234
O3′	−0.2802	−0.2470		−0.2387
P	0.3699	0.5990		0.5043
O7′	−0.4167	0.0016		−0.1631
O6′	−0.3970	−0.1159		−0.1927
O5′	−0.2750	−0.2378		−0.2372
C5′	0.1437	+0.1501		0.1514
H(C5′)	−0.0268	−0.0057		0.0025
H′(C5′)	−0.0395	0.0220		0.0010
N9	−0.0513	0.0359	−0.1663	
C8	0.1568	0.2485	0.1362	
C4	0.2281	0.3166	0.1746	
C5	−0.0346	−0.0522	−0.0410	
N7	−0.2160	−0.0256	−0.3006	
C6	0.3190	0.3762	0.2791	
N1	−0.1110	−0.1063	−0.1275	
C2	0.2118	0.3038	0.1576	
N3	−0.1296	−0.1157	−0.1518	
N6	−0.2180	−0.1743	−0.2339	
H(N6)	0.1603	0.2067	0.1359	
H′(N6)	0.1803	0.2130	0.1616	
H(C8)	0.0419	0.1231	−0.0098	
H(N1)	0.1642	0.2072	0.1367	
H(C2)	0.0480	0.0880	0.0091	
O8′		−0.0752		
O9′		−0.0669		
H(O8′)		0.2307		
H′(O8′)		0.2335		
H(O9′)		0.2298		
H′(O9′)		0.2253		
Cu		−0.9432		

cAMP 分子水解反应的实验研究结果：水解反应是P原子上的亲核取代反应，可以推测金属离子 Cu^{2+}与 cAMP 的环外氧原子的配位相互作用可以促进 cAMP 分子的水解反应的进行。

从江寿平和唐明生等的计算结果看，对于腺嘌呤碱基，金属离子与其配位的最优位置似乎不同，但从图 10-9 中同样可以看出，当 Na^+ 与腺嘌呤的 N_3、N_7 原子单独配位时其相互作用势有极小值（-37.78、-31.20 kcal/mol）。它们都接近于其相互作用势的最小值，因此 Na^+ 与 A 的最优配位模式易受环境影响而可能产生 Na^+与 A 的 N_3、N_7 原子配位的最优配位模式。必须指出，我们这里所讨论的例子仅只是 Na^+ 与单个碱基的最优配位模式。然而，当 Na^+ 与配对碱基（A-T、G-C）或DNA中双联体 $\downarrow {}^{AT}_{GC} \uparrow$、$\downarrow {}^{AT}_{CG} \uparrow$ 相互作用时，它们之间相互作用势的极值情况与单个碱基跟Na^+相互作用势的极值情况是有较大变化的（详见文献[5]）。

总之，尽管金属离子与核酸相互作用的配位位置问题的理论计算还存在着许多困难，但我们仍有可能通过金属离子与核酸中的碱基、碱基对、螺旋双联体或单个核苷酸等相互作用的理论计算，为了解金属离子在核酸发挥其生物功能时所起的作用获得一定的轮廓性认识。

10.3.2 金属离子对核酸构象稳定性的影响

我们知道，核酸的生物学功能是与其一定的构象相联系的。近些年来，核酸构象的研究引起了人们极大的兴趣，国际上许多实验室都开展了这方面的研究。从量子生物学角度研究核酸分子的构象变化对其生物学功能的影响已发展成为十分活跃的领域（参见第 9 章）。在导致核酸构象发生变化的因素中，金属离子的影响是不可忽视的因素，因而有关这方面的量子力学计算也受到了重视。

实验已经证实，由于 DNA 的聚电解质（polyelectrolyte）性质，它的构象的稳定性极大地与所处离子环境有关。随着相反的离子浓度的增加，DNA 的解链温度也增加，这点已为单价的[22]、二价的[23]和三价的[24]金属阳离子的研究所证实。它表明相反离子对双螺旋起着稳定的作用。通过特定的离子类型和离子浓度，已可在DNA中引起二级结构的转变。Pohl和 Jovin[25]已经观察到在交替的共聚物（dG-dC）中由于溶液的离子强度被改变而存在一种可逆的构象变化。Haulon 及其合作者通过园二色谱分析得出 DNA 的 B→C＋A 构象变化的引起是由于离子类型和离子强度在起作用[26]。还有人证明闭合的环状DNA 的双重转动角取决于离子类型和浓度[47]。这样一些以及其它一些实验都证实了，核酸的结构对于与之相作用的金属离子的浓度和类型颇为敏感。

对于DNA中这种专一性结构变化与特定的相反离子的结合方式的理论研究，已经发表了一些报告，其中提出了两种有关相反离子结合的模型：一种是直接的"定点结合"（site binding），这种结合出现在相反离子固定于特定位置，譬如一个磷酸或一个碱基时；另一种假想的模型是"云雾状结合"（atmospheric binding），在这种模型中被结合的相反离子在带正电荷的磷酸基团附近形成可流动的云。这两种模型的普遍性，迄今还没有明确的实验验证，但赞同这一种或另一种模型的实验还是比较多的。例如碱金属与 DNA 相互作用的核磁共振研究表明[28]，^{23}Na 和^{87}Rb 有一种象征性流动结合的驰豫速率，而电泳流动性研究则支持在特定位置结合 Li^+、Na^+和 K^+。

以上我们就金属离子与 DNA 的结合方式的研究做了简略叙述，下面将介绍定点结合方式

表 10-4 某些金属阳离子-$(OH)_2PO_2^-$配合物的平衡几何构型和能量①

M②	$\angle O_1PO_2$	d(PM)	$d(MO_1)$	$\angle PO_1M$	能 量
Li^+	115.55	2.29	1.95	82.2	−649.0516
Li^+	107	2.33	1.88	87.1	−649.0573
Be^{2+}	115.55	1.92	1.69	73.8	−655.0000
Be^{2+}	95	2.00	1.48	85.0	−655.7614
Na^+	115.55	2.66	2.25		
Na^+	110	2.70	2.21		−803.4261
Mg^{2+}	115.55	2.34	1.99		−840.9210
Mg^{2+}	102	2.43	1.89	91.5	−840.9432
LiH_2O^+	107	2.33	1.88	87.1	−725.1289
BeH_2O^{2+}	95	2.00	1.48	85.0	−732.1092

① 所用单位：距离Å、角度、能量Hartree。

② 相同离子的两行中，第一行角O_1PO_2用实验值，距离进行优化；第二行两者都进行优化。金属-H_2O距离未优化。

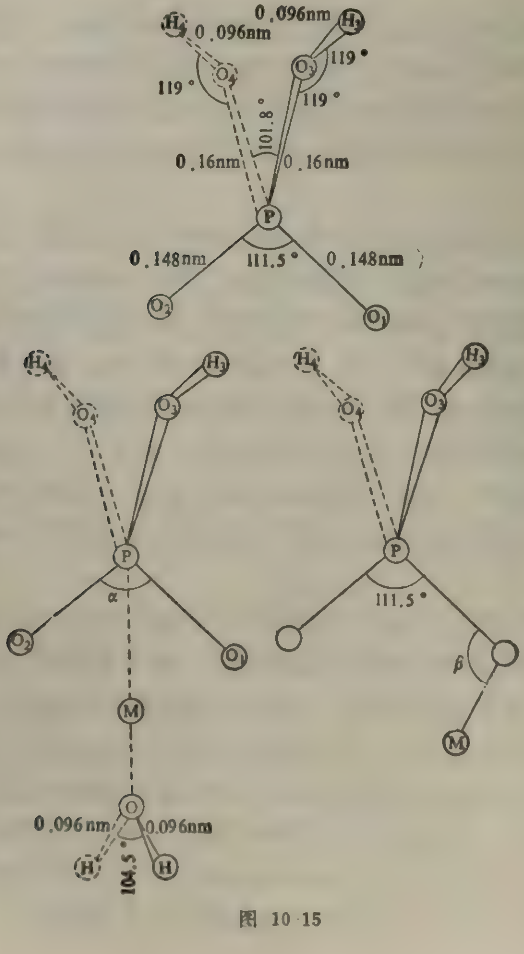


图 10-15

的ab initio SCF 计算以及由 Pack 提出的一种新的理论模型[29]。

Liebmann 等于 1982 年发表了他们采用具有扩展的、双 ξ 正极化基组的 ab initio SCF 法，对$(OH)_2PO_2^-$与 Li^+、Na^+、Be^{2+}、Mg^{2+} 形成的磷酸阴离子-金属阳离子配合物进行研究的论文[30]。对几何构型(图 10-15)选用了在 B-DNA 中固定的 O_1PO_2 键角以及对该键角进行优化处理的两种计算结果列于表 10-4。这个表表明了，在形成配合物时 O_1PO_2 角明显减小，其阳离子影响的次序为 $Na^+<Li^+<Mg^{2+}<Be^{2+}$。

在表 10-5 中，列出了 *A*、*B*、*C* 三种过程的生成热。过程 A 给出的生成热表明配合物的稳定性按 $Na^+<Li^+<Mg^{2+}<Be^{2+}$而增加，$Be^{2+}$形成的配合物差不多比 Na^+稳定 4 倍。过程 B 指出 Li^+和 Be^{2+}与单个水分子的水合热，假定零点能校正忽略，计算的 Li^+的能量与实验值非常接近，从而提供了对这一计算结果的可信性。过程 *C* 给出的则是水合的 Li^+、Be^{2+}与$(OH)_2PO_2^-$之间的配合物的生成热，当金属离子被水合时，值得注意的是，它们与$(OH)_2PO_2^-$ 形成了较弱的配合物但仍保持其相对稳定性。

表 10-5　三种过程的生成热(kcal/mol)

M^{z+}	*A*	*B*		*C*
	$(OH)_2PO_2^-+M^{z+}\to[(OH)_2PO_2M]^{z-1}$	$H_2O+M^{z+}\to(H_2O\cdot M)^{z+}$		$(OH)_2PO_2^-+(H_2O\cdot M)^{z+}\to[(OH)_2PO_2\cdot M\cdot H_2O]^{z-1}$
		计　算	实　验	
Na^+	−138.2			
Li^+	−162.5	−33.3	−34	−145.6
Mg^{2+}	−359.5			
Be^{2+}	−494.7	−143.8		−414.7

对 Li^+、Be^{2+}和 $(OH)_2PO_2^-$的 C_s 对称性配合物计算的总能量有力地提示，金属阳离子与磷酸形成配合物时宁可采用 C_{2v}对称性。

表 10-6 给出了根据对$(OH)_2PO_2^-$、对C_{2v}和C_s进行了优化的(阳离子-$H_2PO_4^-$)以及对水合 Li^+、Be^{2+}与$(OH)_2PO_2^-$的配合物进行了 Mulliken 布居分析而计算的总原子电荷，这些结果可以作为在配合物中从阴离子到阳离子的电子“捐赠”的度量。从此表可以看到，在 C_{2v} 对称性配合物中，从$(OH)_2PO_2^-$到阳离子电子转移的程度分别是 0.05，0.46，0.75 和 1.4 e，即 $Na^+<Li^+<Mg^{2+}<Be^{2+}$。在不稳定的 C_s 对称性配合物中，转移到 Li^+和 Be^{2+}的电荷要少一些。

至于阳离子-水配合物，$(LiH_2O)^+$和$(BeH_2O)^{2+}$，总的电子电荷表明，H_2O 分别“捐赠”0.20e 和0.70 e 给阳离子。在$(H_2PO_4\cdot Li\cdot H_2O)$配合物中，电荷从$(OH)_2PO_2^-$和 H_2O 转移到 Li 的程度大致与具有单个给予体的二元配合物中的情形相同，而在$(H_2PO_4\cdot Be\cdot H_2O)^+$配合物中，当两种给予体都存在时，来自各个给予体的电子转移减少。

从表 10-7 可见，金属-磷酸配合物的共价特征的程度，由 M—O 键的重迭布居分析来看也很明显。对 Na^+、Li^+、Mg^{2+}、Be^{2+}的磷酸配合物其值分别为 0.02，0.35，0.47 和 0.74。该表还表明，磷—阴离子氧(O_1，O_2)键由于金属离子与这些氧的结合而减弱。相反，磷—骨架氧(O_3，O_4)则

表 10-6　金属阳离子-$(OH)\ PO_2^-$配合物中总原子电荷和电荷转移程度*

类　型	对称性	磷	酸	基	团				H_2O			
		q_P	q_{O_1}	q_{O_2}	q_{O_3}	q_H	Σq	q_M	Δe_1	q_O	q_H	Δe_2
$H_2PO_4^-$	C_{2v}	-1.71	−0.97	−0.97	−0.75	+0.35	−1.00					
H_2PO_4Li	C_{2v}	−1.81	−0.88	−0.88	−0.65	+0.36	−0.54	+0.54	0.46			
H_2PO_4Na	C_{2v}	+1.71	−1.01	−1.01	−0.68	+0.36	−0.95	+0.95	0.05			
$(H_2PO_4Be)^+$	C_{2v}	+1.80	−0.52	−0.52	−0.57	+0.39	+0.40	+0.60	1.40			
$(H_2PO_4Mg)^+$	C_{2v}	+1.71	−0.77	−0.77	−0.59	+0.38	−0.25	+1.25	0.75			
H_2PO_4Li	C_s	+1.69	−0.89	−0.99	−0.69	+0.36	−0.85	+0.85	0.15			
$(H_2PO_4Be)^+$	C_s	+1.56	−0.76	−0.78	−0.62	+0.38	+0.16	+0.84	1.16			
H_2O										−0.69	+0.35	
$(LiH_2O)^+$	C_{2v}							+0.80		−0.62	+0.41	0.20
$(H_2PO_4LiH_2O)$	C_{2v}	+1.83	−0.88	−0.88	−0.66	−0.36	−0.53	+0.38	0.46	−0.62	+0.39	0.16
$(BeH_2O)^{2+}$	C_{2v}							+1.30		−0.25	+0.48	0.70
$(H_2PO_4BeH_2O)^+$	C_{2v}	+1.74	−0.56	−0.56	−0.58	+0.38	+0.22	+0.32	1.22	−0.38	+0.42	0.46

*、Δe_1表示磷酸→M，Δe_2表示H_2O→M。

表 10-7　金属阳离子-$(OH)_2PO_2^-$配合物的键重迭布居

类　型	对称性	$P—O_1$	$P—O_2$	$P—O_3$	$O_3—H$	$M—O_1$	$M—O_{H_2O}$	$(OH)_{H_2O}$
$H_2PO_4^-$	C_{2v}	1.22	1.22	0.53	0.73			
H_2PO_4Li	C_{2v}	0.83	0.83	0.73	0.69	0.35		
H_2PO_4Na	C_{2v}	0.99	0.99	0.70	0.70	0.02		
$(H_2PO_4Be)^+$	C_{2v}	0.59	0.59	0.89	0.67	0.74		
$(H_2PO_4Mg)^+$	C_{2v}	0.73	0.73	0.82	0.68	0.47		
H_2PO_4Li	C_s	0.54	1.37	0.67	0.72	0.30		
$(H_2PO_4Be)^+$	C_s	0.21	1.33	0.67	0.70	0.91		
H_2O								0.69
$(LiH_2O)^+$	C_{2v}						0.16	0.65
$(H_2PO_4LiH_2O)$	C_{2v}	0.81	0.81	0.73	0.68	0.36	0.22	0.65
$(BeH_2O)^{2+}$	C_{2v}						0.87	0.63
$(H_2PO_4BeH_2O)^+$	C_{2v}	0.62	0.62	0.99	0.67	0.67	0.87	0.63

由于金属离子与这些氧结合而使键增强，且金属离子水合作用显示出对 Be^{2+}—磷酸键稍有减弱而对 Li^+—磷酸键的影响则很小。

由以上讨论可以看出，在没有直接的实验证据以确定金属离子是否通过改变多聚体组分单元的电荷分布或静电势，以使多核苷酸的构象和稳定性受到影响的情况下，Liebmann等的计算指示了不同的金属离子有不同的作用机理，这对于探讨金属离子对核酸构象和稳定性的影响是有意义的。表 10-6 和表 10-7 还告诉我们，对于 Na^+ 只有非常小的电荷转移至它，而 Na^+—O 键的重迭布居则可忽略。这样的 Na^+在与磷酸的结合中实际上是完全离子型的，仅通过它的势对磷酸势场的迭加来改变核酸的固有稳定性。而Li^+、Be^{2+}、Mg^{2+}则以一定量的电子转移和金属离子—氧键的重迭布居来形成共价复合物。与单价阳离子比较，二价阳离子 Mg^{2+}和 Be^{2+}对于磷

酸电荷分布有更大影响，尤为突出的是 Be^{2+}。根据对离子大小与共价程度关系的分析，可以预料比钠大的碱金属离子也将形成离子型的配合物，而碱土金属离子如象 Ca^{2+}、Sr^{2+} 仍将形成具有可与 Li^+相比拟的电荷转移配合物。

除去在电子结构方面的影响外，金属离子在与磷酸的结合中还影响了特定的 O_1PO_2角，对于Na^+、Li^+、Mg^{2+}、Be^{2+} 而言，该 $\angle O_1PO_2$ 的角度分别为 110°，107°，102°和 95°（见表 10-4），均不同于 DNA 纤维中所观察到的 115.5°。这一次序与在各自的配合物中从磷酸到阳离子电子转移的相对程度相一致，反映出了 O_1—O_2排斥的降低。这可从表 12-6 看出，在表中 Na^+、Li^+、Mg^{2+}、Be^{2+} 配合物的这两个氧原子具有的净原子电荷分别是－1.01，－0.88，－0.77和－0.52。每个氧原子电荷的减少允许它们更为靠近，从而导致角度的减小。然而，处于外部的磷酸几何构型的这种变化将会沿着多聚物的骨架传播，使 O_1PO_2 角关闭，O_3PO_4 角打开，最后引起核酸构象的变化。

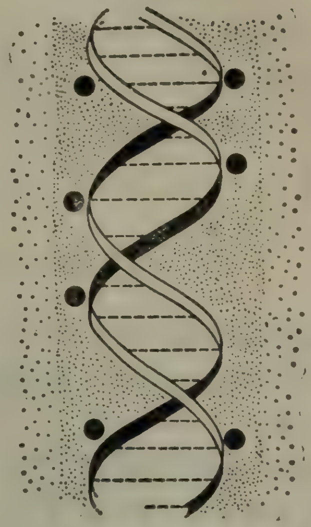


图 10-16　金属阳离子与 DNA 片断结合示意图
大的黑圆点表示固定的相反离子；在多阴离子邻近的小黑点表示可流动的相反离子云雾；远离多阴离子的稀疏小黑点表示低浓度的自由金属离子。

现在让我们进一步比较Li^+ 和 Be^{2+}的单水合作用焓和金属-磷酸配合物的生成焓（见表 10-5），因为这样的比较将有助于我们了解这两种金属与磷酸的相互作用。其中，磷酸-Li^+ 相互作用为 H_2O-Li^+相互作用的 4.4—4.9 倍那么大，而对于 Be^{2+}，其比率是 2.9—3.4。这说明在与磷酸的直接定点结合中，比起 Be^{2+}来，Li^+能够失去更多的水合水。由此可以得出这样的结论：Li^+更容易同磷酸形成直接的配合物，而 Be^{2+} 则能够以保持第一水合层的方式相互作用。

基于 10.3.1 所述的各种不同的实验事实，为了解释金属离子和核酸的相互作用，Pack 提出了一种新的理论模型——高盐模型(higt-salt model)[29]。其图示见图 10-16。

Pack 将 DNA 多聚体看作由实验确定的几何构型上，经理论上确定的电荷分布的一串磷酸基团来考虑。这些磷酸基团的 α 部分(即占用位置)含有固定的金属离子。就每一磷酸-金属离子对而言，其相对位置及电荷分布由 ab initio MO 理论确定。其余的未占用位置$(1-\alpha)$部分有一种相反离子云，通过引入以 Debye-Hückel 近似表达[32]的有效介电常数来屏蔽这些未占用磷酸基团的势。于是，含有 m 个磷酸基团的双螺旋的静电自由能可表示为

$$E_{el}=\sum_{i=1}^{m}\sum_{j>i}^{m}[\alpha^2E_{ij}^{(1)}+\alpha(1-\alpha)E_{ij}^{(2)}+$$

$$+\alpha(1-\alpha)E_{ij}^{(3)}+(1-\alpha)^2E_{ij}^{(4)}] \qquad (10.6)$$

(12.6)式中的四种相互作用类型分别表示：(1) 磷酸基团 i 和 j 都由定点结合的金属离子所占

用：(2) 磷酸 i 占用，磷酸 j 未占用；(3) 磷酸 i 未占用而磷酸 j 占用；(4) 磷酸 i、j 位置都未占用。各类相互作用能表达如下：

$$E_{ij}^{(1)}=\sum_{l=1}^{n}\sum_{l'=1}^{n}\frac{q_l q_{l'}}{|\boldsymbol{r}_{li}-\boldsymbol{r}_{l'j}|} \tag{10.7}$$

$$E_{ij}^{(2)}=\sum_{l=1}^{n}\sum_{l'=1}^{n}\frac{q_l q_{l'}^0}{|\boldsymbol{r}_{li}-\boldsymbol{r}_{l'j}|}De^{k|r_i-r_j|} \tag{10.8}$$

$$E_{ij}^{(3)}=\sum_{l=1}^{n}\sum_{l'=1}^{n}\frac{q_l^0 q_{l'}}{|\boldsymbol{r}_{li}-\boldsymbol{r}_{l'j}|}De^{k|r_i-r_j|} \tag{10.9}$$

$$E_{ij}^{(4)}=\sum_{l=1}^{n}\sum_{l'=1}^{n}\frac{q_l^0 q_{l'}^0}{|\boldsymbol{r}_{li}-\boldsymbol{r}_{l'j}|}De^{k|r_i-r_j|} \tag{10.10}$$

在(10.2)式至(10.5)式中，指标 l 和 l' 分别表示基团 i、j 的原子，q_l 和 q_l^0 分别针对占用了的和未占用的位置而言的原子电荷 r_{li} 是基团 i 和l 原子的位矢，k 是 Debye-Hückel 参数，D 是介电常数。

从这一理论模型出发，Pack 的计算能够解释在与 DNA 的结合当中碱金属间的差异，其结果与碱金属诱导 DNA 由 B 形式向 C 形式转变中所观察到的次序一致。理论计算也与对这些金属离子所观察到的缔合常数以及所观察到的作为离子强度函数的解链温度的变化一致。且由计算结果还可得出这样的结论，即在离子强度低时，所有的碱金属阳离子都将对 DNA 的结构产生类似的影响，结合方式主要是一种在聚阴离子附近的可流动性相反离子的缔合。当碱金属离子浓度增加到 1～2 mol/L 或者更高，便有相当一部分磷酸被固定的相反离子所占用。而这种碱金属离子对某些磷酸位置的占用正好反映出在引起 DNA 发生构象变化中不同离子的不同效应。

以上我们仅就金属离子与核酸间相互作用的量子生物学研究的两个方面作了简述，当然这是远远不够的。但是，我们清楚地看到，通过对金属离子与核酸相互作用的研究，人们已经认识到，核酸的杂环碱基、磷酸基团和核糖(或脱氧核糖)的若干位置均可成为实际上与金属离子相互作用的位置，不同的金属离子有不同的优先结合，同一金属离子因反应条件的变化可以改变其结合的位置。总之，金属离子与核酸间的相互作用是复杂而多样的，对核酸结构和功能的影响同样是复杂而多样的。譬如在核酸传递遗传信息的许多过程中，都离不开金属离子的作用，尽管这些作用的许多细节目前还不十分清楚，但近几年来量子生物学在这些问题上的研究已经获得不少有益的信息。我们的任务之一，就是要更深刻地揭示这些作用的机理，并尽可能地在电子水平上给予解释。

10.4 金属离子与蛋白质的相互作用

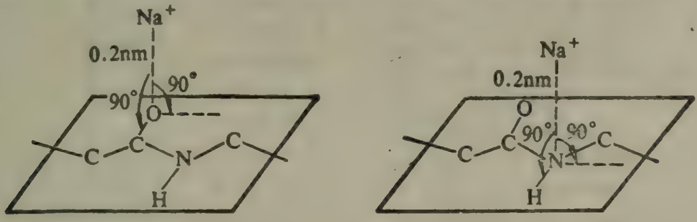
大多数天然存在的金属离子常常与蛋白质结合在一起，因而，研究金属离子与蛋白质之间的相互作用，以及这些相互作用对蛋白质结构和功能的影响，乃是一个极为重要的课题。

由 10.2.1 知，所有的氨基酸、肽和蛋白质均可作为配位体和金属离子相互作用而生成配合物。这些配合物在生物体中的功能因其结构而异。有些金属蛋白质用于金属离子本身的贮存和运送，如铁蛋白用于贮铁，铁传递蛋白用于运送铁，血浆兰铜蛋白则在血浆中和血清中贮存铜等等；有些金属蛋白质则用于作为电子传递体，如细胞色素类在线粒体的氧化磷酸化和叶绿体的光合作用中起作用，铁硫蛋白则在光合作用、固氮作用和甾族化合物的生物合成中起重要作用。还有一些金属蛋白质能够运送氧气供生物体所需，称为氧载体，如血红蛋白、肌红蛋白、血紫蛋白和血清蛋白等。

以下我们将主要讨论金属-蛋白质配合物的量子生物学研究。

10.4.1 金属离子对蛋白质能带结构的影响

蛋白质中能量和电荷迁移不仅在它的生物学功能中起着重要的作用，而且也展示了美好的应用前景(如有机半导体，有机超导体)。为了描述这种能量和电荷迁移，人们就蛋白质中不同的氨基酸残基，这些氨基酸残基侧链的无序排列、蛋白质构象的变化以及水和金属离子等对能量和电荷迁移的影响，进行了若干理论研究。其中，有关蛋白质的能带结构以及金属离子对蛋白质能带结构影响的研究，已经取得了一些进展。我国的科学者肖奕和陈润生等[33]运用最近邻晶胞相互作用近似下的 CNDO/2 SCF LCAO CO(晶体轨道)法研究了金属 Na^+ 对蛋白质能带结构的影响。他们选取具有代表性的周期蛋白质模型——聚甘氨酸[以下简称 poly(Gly)]，对两种 Na^+ 与 poly(Gly)作用的复合体(见图 10-17*A* 和 *B*)作了理论计算。



A *B*

图 10-17 *A*. poly($Gly-Na^+$)I，*B*. poly($Gly-Na^+$)II中的Na^+与肽平面的几何关系

表 10-8(*a*)和(*b*)分别是 poly($Gly-Na^+$)I 和 poly($Gly-Na^+$)II 的能带结构。表 10-9 是它们的价带和导带的位置、带宽的比较。表 10-10 则是它们的禁带宽度的比较。

从表 10-8(*a*)和表 10-8(*b*)可以看出，当 poly(Gly)与 Na^+作用时，整个体系的能带位置向下移动了相当大的距离，几乎所有的能级的能量值都为负。比较 poly($Gly-Na^+$)I 和 poly($Gly-Na^+$)II 不难看出，Na^+ 结合到肽平面不同的原子，其能带结构变化的程度也不相同。当 Na^+ 结合到羰基氧原子时，禁带宽度与 poly(Gly)[34]的(10.421eV)相比，没有改变；价带的宽度增加为 poly(Gly)(0.381eV)的 2 倍，导带宽只有 poly(Gly)(0.490eV)的四分之一。而当 Na^+结合到肽键的氮原子时，禁带宽减小了约 0.9eV，价带宽也是 poly(Gly)的 2 倍，但导带宽变化很小，仅减小了～0.03eV。在这两种情况下，导带都比价带窄。poly($Gly-Na^+$)I和

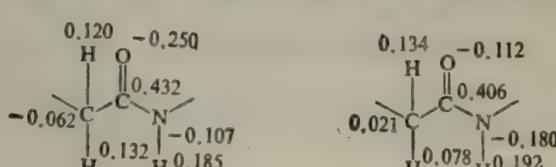
表 10-8 (*a*) poly(Gly-Na⁺)I的能带结构(单位：eV)

带　数	$E_{min}(k.a)$	$E_{max}(k.a)$	ΔE	
28	0.762(0)	2.014(π)	1.252	
27	0.027(π/2)	0.163(0)	0.136	
26	−1.279(3π/4)	−0.762(0)	0.517	
25	−3.102(0)	−2.667(π)	0.435	
24	−4.354(9π/10)	−3.510(0)	0.844	
23	−5.415(π/4)	−4.979(3π/42)	0.436	} 2.395 (带 23–19)
22	−5.689(π/2)	−5.415(π)	0.274	
21	−6.558(0)	−5.605(9π/10)	0.953	
20	−6.884(0)	−6.286(2π/5)	0.598	
19	−7.374(0)	−6.639(π)	0.735	
18	−7.728(0)	−7.428(π)	0.300	
17	−8.462(2π/5)	−8.027(3π/4)	0.435	} 1.252 (带 17–16)
16	−9.279(0)	−8.326(π)	0.953	
15	−9.605(0)	−9.333(3π/5)	0.272	} 0.544 (带 15–14)
14	−9.877(0)	−9.551(3π/5)	0.326	
13	−13.306(0)	−11.211(3π/4)	2.059	
12	−14.884(π/2)	−14.775(0)	0.109	
11	−26.013(0)	−25.305(π)	0.708	
10	−27.727(0)	−27.264(2π/5)	0.463	
9	−29.496(0)	−28.353(π)	1.143	
8	−32.271(π/4)	−31.155(π)	1.113	
7	−35.074(π)	−32.734(0)	2.340	
6	−37.387(π)	−36.434(π/4)	0.953	
5	−40.026(0)	−38.965(2π/5)	1.061	
4	−41.958(π/4)	−40.407(3π/4)	1.551	
3	−51.862(π)	−46.203(0)	4.659	
2	−58.692(π)	−56.733(π/2)	1.959	
1	−72.297(0)	−59.998(π)	12.299	

注：表中E_{min}和E_{max}分别为带的下边和上边值，ΔE为带宽，k为波数，a为平移周期。

poly(Gly-Na⁺)II 导带和价带宽的比分别为～1/7 和 5/7(参见表 10-9 和 10-10)。Ladik 和 Rozsnyai 等[35]对 DNA 能带结构研究结果也表明，金属离子对能带结构的影响和结合的部位有关。

从图 10-18(*A*)和(*B*)可以看出，与 poly(Gly)[34]相比，原子净电荷变化较大的是 H^a和 N 原子，约改变了 0.1a.u.。由于 Na^+ 的吸引，电子发生转移。在 poly (Gly-Na⁺)I 和 poly



A　　　　*B*

图 10-18 *A*. poly(Gly-Na⁺)I, *B*.poly(Gly-Na⁺)II 的原子净电荷(单位：a.u.)

表 10-8(b)　poly(Gly-Na⁺)II的能带结构(单位：eV)

带　数	$E_{min}(k.a)$	$E_{max}(k.a)$	ΔE
28	0.054(0)	1.143(π)	1.089
27	−0.789(π/4)	−0.463(π)	0.326
26	−1.660(3π/5)	−1.497(0)	0.163
25	−3.837(0)	−3.537(π)	0.300
24	−4.843(π/2)	−4.272(π)	0.571
23	−6.041(0)	−5.605(π)	0.436
22	−6.694(0)	−6.422(π/2)	0.272
21	−7.374(π)	−6.694(0)	0.680
20	−8.136(0)	−7.592(9π/10)	0.534 } 1.170
19	−8.762(0)	−8.136(π)	0.626
18	−9.034(π/4)	−8.462(π)	0.572 } 0.844
17	−9.306(0)	−9.034(3π/4)	0.272
16	−10.0040(0)	−9.605(π)	0.435
15	−10.503(0)	−10.068(3π/4)	0.435 } 1.115
14	−11.183(0)	−10.394(3π/4)	0.789
13	−14.176(0)	−11.401(π)	2.765
12	−16.516(π)	−16.054(0)	0.462
11	−26.747(0)	−26.040(π)	0.707
10	−28.707(0)	−28.026(3π/5)	0.681
9	−30.611(0)	−29.169(π)	1.442
8	−33.196(π/4)	−32.108(π)	1.088
7	−36.108(π)	−33.523(0)	2.585
6	−38.339(π)	−37.332(2π/5)	1.007
5	−41.441(0)	−40.244(π)	1.197
4	−43.209(π/4)	−41.577(9π/10)	1.632
3	−53.522(π)	−47.019(0)	6.503
2	−59.998(π)	−57.440(π/2)	2.558
1	−74.474(0)	−61.495(π)	12.979

表 10-9　poly(Gly-Na⁺)I, poly(Gly-Na⁺)II的价带和导带(分别用n^*和n^*+1表示)
(单位：eV)

系　统	带　型	E_{min}	E_{max}	ΔE
poly(Gly—Na⁺)I	n^*+1	−14.884	−14.775	0.109
	n^*	−26.013	−25.305	0.708
poly(Gly—Na⁺)II	n^*+1	−16.516	−16.054	0.462
	n^*	−26.747	−26.040	0.707
poly(Gly—H_2O)I	n^*+1	−0.735	−0.272	0.463
	n^*	−11.428	−11.129	0.299
poly(Gly—H_2O)II	n^*+1	−0.544	−0.000	0.544
	n^*	−11.238	−10.612	0.626

表 10-10 poly(Gly—Na^+)I, poly(Gly—Na^+)II的禁带宽度(单位：eV)

系统	禁带宽度
poly(Gly—Na^+)I	10.421
poly(Gly—Na^+)II	9.524

(Gly-Na^+)II 中分别有 0.30e 和 0.45e 电荷转移到 Na^+离子上。后者转移电荷较多，可能是由于在 N 原子上有一孤对电子引起的。在 poly(Gly-Na^+)II 中两个氨基酸残基几乎就有一个电子转移到 Na^+ 离子上。因此，在 poly(Gly)链中两个氨基酸残基就可能在价带中产生空穴。由于价带比较宽，是比较利于空穴迁移的。

综上所述，Na^+ 与 poly(Gly)作用在不同程度上使禁带变窄，使蛋白质向半导体的范围变化。同时，Na^+ 与 poly(Gly)作用大都使价带增宽，且价带比导带宽。这就表明蛋白质分子中的空穴有较大的迁移率。这有利于能量和电荷在蛋白质分子中的迁移。

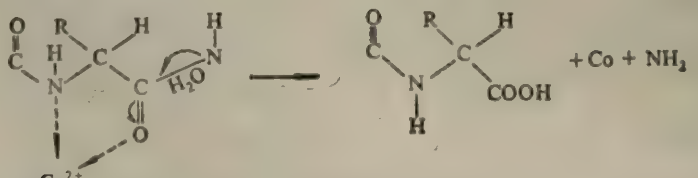
10.4.2 金属的螯合性质对酶活性的影响

大量的研究工作证明，某些含金属的酶中，金属离子在活性中心中占有一定的地位。目前在已知酶中约有四分之一含有紧密结合的金属离子或在催化时须要用到金属离子。在酶学中很多酶机制的解释都要利用金属的螯合性质。这些金属的功能可用各种物理方法加以研究，例如X射线结晶学，核磁共振，电子自旋共振等……，近些年来，随着若干酶和蛋白质的三维结构在原子水平上的阐明，有关金属离子结合酶活性中心的量子生物学研究也活跃起来。但是，正如我们在第一和第二章中所述的那样，困难和潜在的危险是不可低估的。迄今为止，这方面的研究尚处于起始阶段，理论计算也比较粗糙，因而我们仅作简略介绍。

业已知道，酶分子中金属的作用一般是维持酶的一种活性构象，但不少酶还能同时形成由酶通至底物的一种金属桥，如丙酮酸激酶中：

丙酮酸激酶—┬—M
　　　　　　├—ATP
　　　　　　└—丙酮酸

可见除了稳定构象之外，丙酮酸激酶中的金属离子还能将一个底物(ATP)置于适当的位置使它激活。又如，不少肽酶须要 Co^{2+}才能表现最大活性。Co^{2+}的作用就是把酶与底物连接起来，通过与底物肽键中电子供体氧原子和氮原子形成五元环复合物，Co^{2+} 吸收电子而使肽键易被水解：



从 X-射线结晶学和肽链顺序数据可知，在羧肽酶 A、细胞色素 c 等酶中的组氨酸残基与金属的连接和这些酶的活性中心有关。这里的金属很可能通过金属桥配合物作为底物与酶活性中心联系的桥梁，这类桥梁作用主要由于金属离子能吸引电子而使化学键减弱甚至断裂。金属离子这方面的功能比质子强得多，因为一般它带有的正电荷大于质子。而且在中性水溶液中金属离子的浓度，可远远高于水合氢离子的浓度，因此它们能很好地作为酶活性中心通至底物的桥梁。例如碳酸酐酶主要是催化二氧化碳与水生成碳酸。此酶含有锌离子，锌离子能与底物的氧原子形成配位键，构成配合物。由于二氧化碳中的氧原子可作为电子供体，它与锌离子之间自然地形成一种电子移动，这样就使底物中的双键易于断裂而与水分子中的 OH^-结合生成碳酸。

以下，我们以碳酸酐酶活性部位束缚底物的量子化学研究为例，讨论金属离子对酶活性的影响。

碳酸酐酶是被人们发现的第一个含锌的金属酶。多年来，虽然对碳酸酐酶的作用机制作过很多研究，但仍旧所知甚少。然而，由于这种酶的两种分子形态的晶体结构已被测定，[36,37] 因而为运用量子力学方法来研究其催化作用机制提供了可能性。

在活细胞中，碳酸酐酶催化如下反应：

$$H_2O + CO_2 \rightleftharpoons H_2CO_3 \qquad (a)$$

或

$$OH^- + CO_2 \rightleftharpoons HCO_3^- \qquad (b)$$

除了这样的反应外，碳酸酐酶还能催化某些脂肪的水解。

碳酸酐酶中的锌离子可被钴、锰、镍、铜、铅、铁、铍、镁和钙等离子取代。当用钴代替锌时，酶保留了高活性；而用镍和锰取代锌时，则降低了酶的活性；在其它取代情况下，酶的活性消失。

在碳酸酐酶的大分子中有很深的裂缝，锌离子就束缚在裂缝的底部。在这里，锌离子可能是通过咪唑基团与三个组氨酸根相连，而第 4 个配位点看来是被水分子或 OH^-所占据。

Jonsson 等[38,39]用 ab initio 法计算碳酸酐酶催化 CO_2的水合反应表明，在气相时，上述反应(a)的能垒为 52.3kcal/mol，(b)反应则没有势垒。实验研究表明[40,41]，在溶液中，相应的实验活化能分别为 17.7和 13.1 kcal/mol。这就说明，分子环境对于这样的催化反应具有显著影响。实际上，碳酸酐酶活性强烈地依赖于其分子的形态以及金属离子束缚在其活性部位的种类[42]。表 10-11 表示分子环境对 CO_2 水合反应的影响。

Sawaryn 等[43]的理论计算也清楚地说明，在气相时，OH^- 直接进攻 CO_2。在液相时，这种反应的实验测定能垒[41]可以说明，更复杂的机理可能包括另一个水分子强烈地束缚OH^-。由此可见，碳酸酐酶催化 CO_2 的水合反应应可能是由于 CO_2与 OH^- 的直接反应结果，也可能是 CO_2 与水分子反应并在离子化作用中失去质子的结果。但似乎前者更可能些。

Pullman 等[44]在运用赝势从头计算法(参见 2.2.3)计算模型配合物($Zn^{2+}(OH_2)_n$、$Zn^{2+}(OH)_{n-1}(OH^-)$其中 $n=1.2$，4,6；$Zn^{2+}(NH_3)_{n-1}(OH_2)$ 和 $Zn^{2+}(NH_3)_{n-1}(OH^-)$ 其中 $n=$ 2,4 以及三个配合物 $Zn^{2+}(NH_3)_2(ImH)(OH_2)$*、$Zn^{2+}(NH_3)_2(ImH)(OH^-)$和 $Zn^{2+}(NH_3)_2$

* Pullman 等建立的活性部位模型，其中 ImH 为咪唑。

表 10-11　分子环境对CO_2水合反应的影响

		酶　形　态	Δ(kcal/mol)*
图 10-19*A*	B	苏氨酸 199	−16.8
	B	组氨酸 200	a
	C	苏氨酸 199	a
	C	苏氨酸 200	a
图 10-19*B*	B	苏氨酸 199	75.9
	B	组氨酸 200	144.4
	C	苏氨酸 199	−35.3
	C	苏氨酸 200	52.1
图 10-19*C*	B	苏氨酸 199	−50.5
	C	苏氨酸 199	−103.0
	C	*b*	−112.8
	C	*c*	−25.5
	C	*d*	−14.5

* $\Delta=(E_e-E_g)$过渡配合物$-(E_e-E_g)$底物，该式反映分子环境的相对催化活性。E_e表示酶环境的总能量，E_g相应于在气相时的能量。

a．由于几何原因，这种位置不可能出现。

b．仅考虑Me^{2+}和OH^-的影响。

c．仅考虑OH^-的影响。

d．同*c*，但在超分子近似范围内。

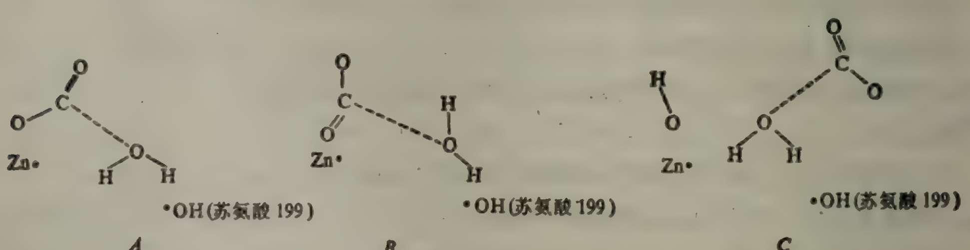


图 10-19　碳酸酐酶活性中心过渡配合物

$(Im^-)(OH_2)$。$n=6$ 和 $n=4$ 分别相当于八面体和四面体几何构型)的配位性质及 CO_2 的反应的基础上，归纳出了若干有意义的结果，例如：

（1）在四配位的配合物中，OH_2 的平均结合能的范围从−78.7——85.9 kcal/mol。当 $n=6$ 时，相应的能量为−69.7——69.9kcal/mol，$n=2$ 时为−102.4——100.7kcal/mol。类似地，OH^-的质子亲合性(p*A*)跟配位数也是明显相关的。譬如由 $Zn^{++}(OH_2)_3(OH^-)$的p*A*值等于 155.6kcal/mol，$Zn^{++}(NH_3)_3(OH^-)$的 p*A* 值等于 158.1kcal/mol，可见在氨取代水后，p*A* 值的改变并不明显，而当咪唑取代氨后，p*A* 值有稍大的改变(对于$Zn^{++}(NH_3)_2(ImH)(OH^-)$，p*A* 等于 177.0kcal/mol)。但是，这些取代对 p*A* 值改变的影响，远不如配位数增加到 6 时(对于 $Zn^{++}(OH_2)_5(OH^-)$，p*A* 值等于 193.8kcal/mol) 那样明显。对配位数为 6 的锌配合物所作的研究表明，配位数增加，结合的水的去质子化能也增加(因而其 pK_a 也将

增加)；结合的 OH^-上的电荷增加约-1.0（于是其亲核性也将增加)。(参见表 10-12)。

(2) 关于活性部位分子模型吸引 CO_2：在 CO_2 的 C 原子和配位的 OH^- 的 O 原子间有明显的吸引相互作用，但跟配位的 OH_2 则没有那样明显的吸引相互作用。

表 10-12 锌配合物的几何构型和能量*

配合物	BE	pA	$\left.\begin{matrix}BE_L\\R_L\end{matrix}\right\}L=OH_2$	OH^-	NH_3	ImH	Im^-
$Zn^{++}OH_2$	−113.6		−113.6				
			1.90				
$Zn^{++}(OH_2)_2$	−204.8		−102.4				
			1.91				
$Zn^{++}(OH_2)_4$	−343.4		−85.9				
			1.98				
$Zn^{++}(OH_2)_6$	−417.9		−69.7				
			2.10				
$Zn^{++}OH^-$	−413.0	81.7		−413.0			
				1.70			
$Zn^{++}(OH_2)(OH^-)$	−474.6	109.7	−102.4	−372.2			
			1.91	1.72			
$Zn^{++}(OH_2)_3(OH^-)$	−567.3	155.6	−85.5	−310.8			
			1.98	1.80			
$Zn^{++}(OH_2)_5(OH^-)$	−603.5	193.8	−69.9	−254.1			
			2.10	1.90			
$Zn^{++}NH_3$	−134.5				−134.5		
					1.98		
$Zn^{++}(NH_3)(OH_2)$	−219.9		−100.7		−119.2		
			1.93		2.0		
$Zn^{++}(NH_3)_3(OH_2)$	−366.4		−80.5		−95.3		
			2.0		2.1		
$Zn^{++}(NH_3)(OH^-)$	−487.4	112.0		−367.7	−119.7		
				1.75	2.0		
$Zn^{++}(NH_3)_3(OH^-)$	−587.8	158.1		−297.3	−96.8		
				1.80	2.1		
$Zn^{++}ImH$	−171.7					−171.7	
						1.87	
$Zn^{++}Im^-$	−367.8	167.9					−367.8
							1.82
$Zn^{++}(NH_3)_2(ImH)(OH_2)$	−384.0		−78.7		−93.2	−118.9	
			2.0		2.1	1.95	
$Zn^{++}(NH_3)_2(ImH)(OH^-)$	−592.5	171.0		−286.6	−93.4	−119.2	
				1.80	2.1	1.95	
$Zn^{++}(NH_3)_2(Im^-)(OH_2)$	−540.6	207.3	−81.8		−96.9		−2[illegible]5.0
			2.0		2.1		1.95

* 距离(Å)，能量(kcal/mol)。

BE=配合物的结合能$=E$(配合物)$-E(Zn^{++})-E(L_1)-E(L_2)\cdots-E(L_n)$。

pA=配合物的质子亲合性[$pA(OH^-)=379.4$；$pA[IM^-)=363.9$]。

BE_L=配合物中配位体L的平均结合能。

R_L=配位体L的O或N原子与τ_n^{2+}间的键距离。

(3) 由表 12-13 可见，咪唑和氨几乎转移相同数量的电荷给 Zn^{2+},这是由于两者对 Zn^{2+} 有几乎相同的引力(有相同的核间距)。这也跟咪唑和氨的静电势具有相似的极小值(其中咪唑为－82.0，氨为－87.7kcal/mol)[44] 的事实是相一致的。

(4) 金属的性质和电荷以及配位体的性质和数目，对于决定结合的 OH^- 的亲核性有明显的影响。

(5) $Zn^{++}(OH_2)_5(OH^-)$有最大的结合能(BE＝－603.5 kcal/mol)。$Zn^{++}(OH_2)_6$的 BE 值(－417.9kcal/mol)比最佳活性部位模型——$Zn^{++}(NH_3)_2(ImH)(OH_2)$的 BE 值(－384.0 kcal/mol)大(见表 10-12)。

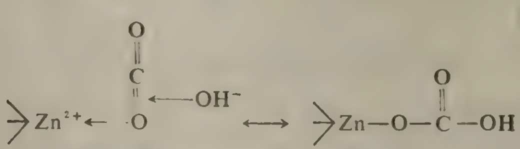
表 10-13　氨和咪唑结合Zn^{2+}的比较*

配　合　物	R	BE	E_C	E_{ex}	E_{del}	电荷转移
$Zn^{2+}\cdots ImH$	1.87	－172.0	－149.5	69.7	－92.2	394
$Zn^{2+}\cdots NH_3$	1.965	－133.9	－127.0	48.9	－55.8	385
	1.87	－132.3	－152.7	77.3	－56.9	382

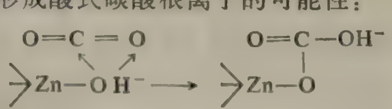
* BE＝配合物的结合能＝$E(Zn^{2+}\cdots L)-E(Zn^{2+})-E(L)$。
E_C＝Coulomb能贡献：未微扰片断Zn^{2+}和 L 的核和电子云间的Coulomb相互作用。
E_{ex}＝交换排斥：Zn^{2+}和 L 的电子之间的Pauli原理排斥。
E_{pcl}＝离域能(电荷转移＋极化)。
距离(Å)，能量(kcal/mol)，电荷转移(毫电子单位)

(6) 锌-结合水的电离导致 CO_2 的重取向，而有利于亲核进攻。

ab initio 分子轨道计算明显地向我们表明，锌离子在很大程度上促进了水的电离。从反应机理看[2]，最可能的是协同反应，即两个试剂从两个方面(锌离子和氢氧离子)同时进攻 CO_2 分子：



在催化中心还存在另一种形成酸式碳酸根离子的可能性：



到目前为止，还不清楚哪一种机理更合适。

10.5　金属-卟啉配合物的分子轨道研究

正如我们已经知道的那样，卟啉被看作是生命的关键分子[45]。而未结合的卟啉并不表现出任何直接的生物学功能[46]。但是，当它们一旦跟金属离子结合成金属卟啉配合物时，就会显示出

极其重要的生物学功能。不仅如此，卟啉型结构跟其它金属离子(如 Fe^{2+}、Mg^{2+}、Cu^{2+} 等)生成的配合物 在生命过程中起着极其重要的作用。

10.5.1 卟啉配合物的分子轨道计算(一)

Pullman 等早在 60 年代初，就运用简单的 Hückel 法对金属离子-卟啉配合物作了分子轨道计算[47]。

在处理金属-卟啉配合物的分子轨道计算时，为了达到简化计算的目的，往往以未取代的卟吩环来代替卟啉。因为卟啉的基本骨架是卟吩(见图 10-20)。当卟吩环的 C 1 至 C 8 上的氢原子，部分或全部被其它基团取代后所得的衍生物即称为卟啉。生物体中常见的卟啉有 8 种，其中以原卟啉 IX(见图 10-21)为最重要，它是血红蛋白、肌红蛋白和细胞色素 c 中的辅基的组成部分。其中，血红蛋白和肌红蛋白 都是血红素蛋白，两者密切相关，它们都含有血红 素基团(作辅基)，而血红素则为 Fe^{2+} 与原卟啉 IX 的配合物。

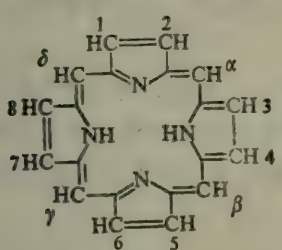
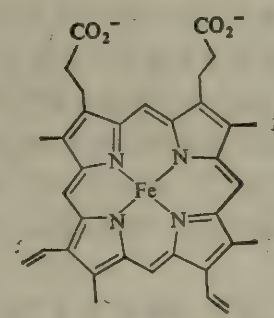
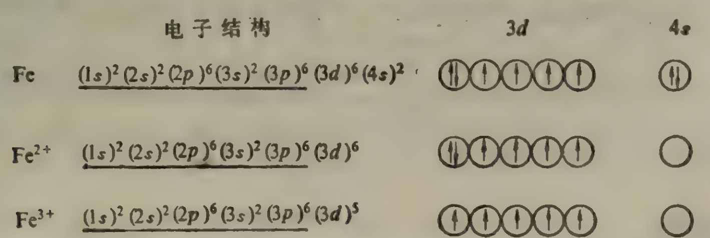


图 10-20 卟吩


图 10-21 原卟啉 IX

由于铁是生物体内含量最多的微量金属元素。尤其在哺乳 动物体内 约有 70%的 铁是以卟啉配合物的形式存在的(其中包括血红蛋白、肌红蛋白、过氧化氢酶以及细胞色素 c 等)，故在这一节里，我们将着重讨论铁卟啉配合物的分子轨道计算。首先，看看铁和它的离子状态的电子结构，然后再讨论铁卟啉配合物的分子轨道计算。中性铁有 26 个电子(如图 10-22)，其中 18 个是处在闭壳状态，余下的 8 个中的2个在4s 轨道，6个分布在 3 d 的 5 个轨道上。在铁卟啉配合物中的铁，有以二价 Fe^{2+} 和三价 Fe^{3+} 形式存在，其电子状态也列在图 10-22 中[48]。


图 10-22 Fe 和它的离子状态的电子结构

划线部分是闭壳结构，可以不予考虑。右侧只表示余下的价电子，↑不成对自旋，↑↓自旋为 0 的状态

10.5.1.1 铁卟吩分子轨道的构造

卟吩分子是一平面共轭结构。如前所述，在 1—8 位置上可有各种取代基。自然界中八种类型的卟吩，其区别在于八个位置上的取代基不同。其中最重要的是原卟啉，它有这样一些取代基：1,3,5,8-甲基，2,4-乙烯，6,7-丙酸基。

在 NH_2 基团中的两个氢原子可以被金属离子置换，其中有 Fe^{2+}、Mg^{2+}、Cu^{2+} 等。卟啉和蛋白质可以通过不同的官能团相联结。

量子化学研究告诉我们，铁与卟吩之间电子的相互作用取决于两者分子轨道的对称性。

卟吩结构属于 D_{4h} 对称群，因而卟吩的 σ 和 π 分子轨道属于群的不同的不可约表示，即 A_{1g}、A_{1u}、A_{2g}、A_{2u}、B_{1g}、B_{1u}、B_{2g}、B_{2u}、E_{1g} 和 E_{1u}。

1. π 轨道　卟吩的 π 轨道属于 A_{1u}、A_{2u}、B_{1g}、B_{2g} 和 E_{1g} 表示。铁阳离子的 π 轨道可以跟属于同一表示的卟吩的 π 轨道相互作用。满足这种对称性条件的是 $3d_{xz}$、$3d_{yz}(E_{1g})$ 和 $4p_z(A_{2u})$。然而，

(1) 在离子型配合物（亚铁或正铁）中，铁的 d 轨道不参与相互作用，跟卟啉的轨道发生相互作用的只是自由金属原子的空的 $4p_z$ 轨道。在亚铁配合物中，铁保持一对 $3d$ 孤对电子和 4 个未配对 $3d$ 电子；在正铁配合物中，铁保持着 5 个未配对电子。

(2) 在共价亚铁配合物中，铁贡献 4 个 $3d$ 电子给配合物的 π 系统，这 4 个 $3d$ 电子位于 $3d_{xz}$ 和 $3d_{yz}$ 轨道。实际上，在这种配合物中 6 个共用的 $3d$ 电子将成对地占据三个最低轨道：$3d_{xy}$、$3d_{xz}$ 和 $3d_{yz}$。在系统的 π 相互作用中，适合于对称性的是 4 个电子占据着的 $3d_{xz}$ 和 $3d_{yz}$ 轨道。

在共价正铁配合物中，阳离子仅贡献 3 个 $3d$ 电子给配合物的 π 系统，这 3 个 $3d$ 电子同样位于 $3d_{xz}$ 和 $3d_{yz}$ 轨道。

卟吩的 π 系统包含 24 个分子轨道和 26 个电子。亚铁或正铁离子配合物是 25 个分子轨道和 26 个电子的系统；亚铁共价配合物是 27 个分子轨道和 30 个电子的系统；亚铁共价配合物则是一个 27 个分子轨道和 29 个电子的系统。

2. σ 轨道　对于 σ 结合，一般只考虑阳离子与卟啉的 4 个中心氮原子间相互作用的计算，而能够跟氮原子结合的铁的轨道是：$3d_{z^2}$ 和 $4s$ 轨道 (A_{1g})、$3d_{x^2-y^2}$ 轨道(B_{1g})，$4p_x$ 和 $4p_z$ 轨道(E_{1u})。

在离子型配合物（亚铁或正铁）中，其余的 $3d$ 轨道未涉及。涉及到的只是跟氮的 σ 轨道结合的阳离子的 $4s$，$4p_x$ 和 $4p_z$ 空轨道。在此情况下，σ 系统仅由 4 个氮原子提供的 8 个电子，组成 7 个 σ 分子轨道。

在共价配合物（亚铁或正铁）中，$3d_{z^2}$ 和 $3d_{x^2-y^2}$ 轨道也跟氮原子的 σ 轨道结合。但铁原子并没有再贡献任何电子来成键。在此情况下，是 4 个氮原子提供 8 个电子，组合成 9 个 σ 分子轨道。

通过选择适当的参数，对一些原子的库仑积分和交换积分进行计算早已实现。有关这些参数的选择和计算，读者可参看文献[49-51]。

10.5.1.2 能级的排布

表 10-14　分子轨道能量(β)及它们的类型

	卟　吩	离子型亚铁和正铁卟啉	共价亚铁卟啉	共价正铁卟啉
			−5.72 σ	−5.72 σ
			−5.05 π	−5.05 π
		−5.05 π	−4.32 σ	−4.32 σ
		−4.50 σ	−4.32 σ	−4.32 σ
		−4.32 σ	−3.98 π	−3.98 π
		−4.32 σ	−3.98 π	−3.98 π
	−3.91 π	−3.91 π	−3.91 π	−3.91 π
	−3.88 π	−3.88 π	−3.15 σ	−3.13 σ
空分子轨道	−3.88 π	−3.88 π	−2.86 π	−2.86 π
	−3.83 π	−2.86 π	−2.68 π	−2.68 π
	−2.68 π	−2.68 π	−2.54 π	−2.54 π
	−2.23 π	−2.23 π	−2.54 π	−2.54 π
	−2.23 π	−2.23 π	−1.92 π	−1.92 π
	−1.92 π	−1.92 π	−0.81 π	−0.78 π
	−0.73 π	−0.73 π	−0.81 π	−0.78 π
	−0.25 π	−0.25 π	−0.73 π	−0.73 π
	−0.25 π	−0.25 π	−0.68 σ	−0.64 σ
	+0.31 π	+0.33 π	−0.08 π	−0.07* π
	+0.54 π	+0.54 π	−0.08 π	−0.07* π
	+0.77 π	+0.77 π	+0.34 π	+0.34 π
	+0.77 π	+0.77 π	+0.54 π	+0.54 π
	+0.77 π	+0.77 π	+0.77 π	+0.77 π
	+0.77 π	+0.78 π	+0.78 π	+0.78 π
	+0.93 π	+0.93 π	+0.87 π	+0.87 π
	+0.93 π	+0.93 π	+0.87 π	+0.87 π
	+1.16 π	+1.16 π	+0.97 π	+0.97 π
占据分子轨道	+1.49 π	+1.49 π	+0.97 π	+0.97 π
	+1.57 π	+1.57 π	+1.16 π	+1.16 π
	+1.57 π	+1.57 π	+1.49 π	+1.49 π
	+1.62 π	+1.62 π	+1.62 π	+1.62 π
		+1.90 σ	+1.62 π	+1.62 π
		+1.96 σ	+1.62 π	+1.62 π
		+1.96 σ	+1.96 σ	+1.96 σ
		+1.96 σ	+1.96 σ	+1.96 σ
			+2.09 σ	+2.09 σ
			+2.24 σ	+2.26 σ

* 专指含三个电子的 π 轨道。

表 10-14 列出了通过量子化学计算 得到的卟吩环和不同类型的铁-卟啉配合 物的分子轨道能量。

当然，我们感兴趣的乃是 π 轨道，尤其是 HOMO 和 LUMO 的能量值。从表 10-14 可见，就这两个轨道的能量而言，离子型配合物与共价配合物之间存在着明显的差别。

由于卟吩本身的 HOMO 和 LUMO 能量的绝对值较小，故卟吩应当是较好的 π 电子给予体和 π 电子接受体。从表 10-14 中我们还可以看出，实际上离子型亚铁和正铁-卟啉配合物的 HOMO 和 LUMO 能量完全相同。不仅如此，就 LUMO 能量而言，卟吩与这两者相比也是完全相同的（－0.25），而卟吩的 HOMO 只是较两者的绝对值低 0.02。这就意味着，离子型亚铁和正铁卟啉既是较好的 π 电子给予体又是较好的 π 电子接受体。理论研究结果还告诉我们，这些 π 电子实际上是卟吩骨架上的那些 π 电子，而铁的 $3d$ 电子并未参与成键，且正是这些 $3d$ 电子，它们的轨道位于成键 π 轨道之上，从而参与了这些配合物的氧化-还原过程。

生物化学的研究表明，包含铁卟啉结构的酶类（如细胞色素、过氧化物酶和过氧化氢酶等）、血红蛋白、肌红蛋白等等都是以一个相同的铁卟啉环系作为辅基或结合铁的化合物。然而，象铁卟啉酶类这样的结合蛋白质，其主要生物化学反应的功能就是集中在铁原子上的。由此可见，理论计算结果（即使是 Hückel 分子轨道法的计算结果）与生物化学的研究结果是符合得较好的。

除以上所述外，从表 12-14 同样可以清楚地看到，HOMO 和 LUMO 能量在共价配合物中是不同的。在这些配合物中，π 轨道包含了卟吩的 p_z 电子和铁原子的 $3d$ 电子，且 HOMO 和 LUMO 的能量分布在某种意义上完全不同于离子型配合物。在共价亚铁卟啉中，HOMO 表现出双重简并，且带有 4 个处于很高能级的电子。这些轨道的存在意味着配合物将具有极强的电子给予体性质。当然，共价亚铁卟啉配合物不具有电子接受体性质。

10.5.1.3 电荷分布

图 10-23 给出了 Pullman 等[47]计算得到的 π 电子分布图：

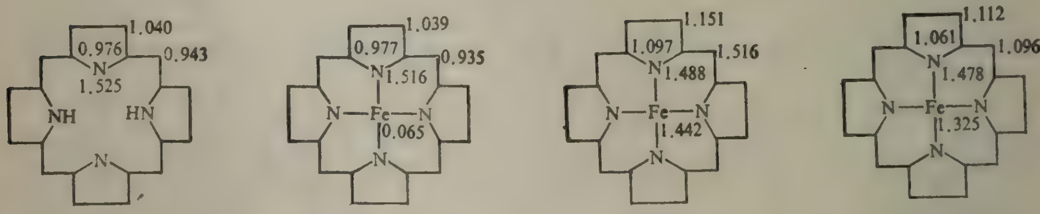


图 10-23 π 电子的分布

从此图可知，在离子型配合物（亚铁或正铁）中，铁原子获得卟啉 π 电子系统的很小一部分电荷；在共价配合物中，铁原子却把大部分 $\pi(d)$ 电子给予卟啉环。转移的部分在亚铁配合物中是 2.558 e，在正铁配合物中是 1.675 e。不难看出，三价阳离子是作为一个较二价阳离子为弱的 d_ϵ-电子给予体而起作用的。

不言而喻，除 π 电子外，还必须考虑 σ 电子在分子中心部分的分布，其分布的特征是电荷的强烈位移（见图 10-24）。在配合物的各类型中，σ 电荷的位移是从卟啉环的氮原子向铁阳离子方向进行的。被阳离子获得 σ 电荷的量在各配合物中不同。已经知道，在共价配合物中位移的数值比离子型配合物大得多，而共价正铁卟啉又比共价亚铁卟啉大些。如果将图 10-23 和图 10-24的电荷分布迭加就可得到总的电荷分布。在离子型配合物中，阳离子所带的总电荷为0.520e；在共价亚铁配合物中，阳离子所带的总电荷为 2.472 e，而共价正铁配合物中为 2.410 e。对铁



图 10-24 σ电子在配合物中心部分的分布

而言,在离子型亚铁卟啉中,其净电荷是 1.480 e,而在离子型正铁卟啉中是2.480 e;在共价亚铁卟啉中为 3.528 e,在共价正铁卟啉则为 3.590 e。按照这些计算结果不难推测,由于细胞色素的亚铁阳离子存在着较高的净正电荷,因此将此血红蛋白的亚铁阳离子更显电负性。同时,我们也注意到,共价配合物的外围碳原子上电荷的分布超过了离子型配合物。

另一方面,电荷分布的研究将有利于说明诸如细胞色素的氧化-还原跟分子轨道形成机制的关系等等这样一些问题。图 10-25 显示共价配合物最高占据轨道的电子分布。看来,即使铁离子有最大的局部浓度(这一部分只占电子数的 20/100),剩下的电子被分布在卟啉的 16 个碳原子的范围。

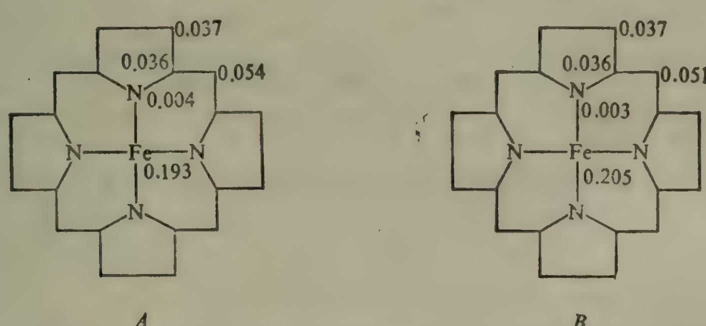


图 10-25 共价亚铁卟啉(*A*)和正铁卟啉(*B*)最高占据轨道的电子分布

10.5.2 卟啉配合物的分子轨道计算(二)

在 10.5.1 中,我们介绍了 Pullman 等对于卟啉配合物分子轨道计算的早期工作。在那以后,半经验分子轨道法乃至 ab initic法等一些更精确的量子化学方法也相继被用于生物金属配合物的分子轨道计算。随着这些计算方法和理论模型的进展,有关金属卟啉电子结构的理论研究正在深入发展。近些年来,已经发表了大量的研究报告,其中仅仅涉及铁卟啉电子结构的报导就有不少,如文献[52—62]。因此另辟一节,以便重点讨论 ab initio 计算。

事实上,早在 1974 年,Almlöf 就用最小基集合的ab initio分子轨道法对卟吩作了理论计算[52]。1981 年,Kashiwagi又用包括 300个收缩高斯型轨道(CGTO)的双重Zeta基集合ab initio LCAO SCF MO 法对平面的 Fe-卟吩作了计算[53]。以下我们将着重讨论这一理论计算的一

些结果，以便与 12.5.1 所述 Pullman 等的工作作某些比较。如前所述，铁卟吩被看作是具有 D_{4h} 对称性的平面分子。Kashiwagi计算的分子模型如图 10-26 所示，它包括 186 或 185 个电子。计算结果，可归纳为：

1. 总能量和电荷分布

表 10-15 和表 10-16 分别列出了 d 轨道的形式占据数和总能量；六个低能态的总原子轨道集居和 Fe-N 的重迭集居。

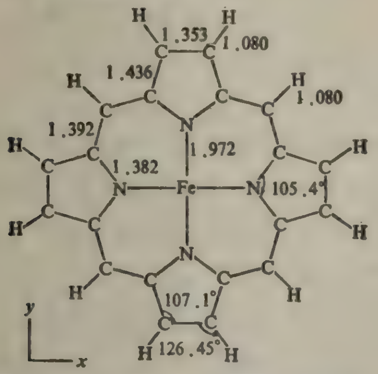


图 10-26 Fe 卟吩 D_{4h} 的假定几何图形

从这两个表中可以看出，对中性Fe-卟吩的9个低能态和5个电离态的SCF波函数所作的计算（参见表 10-15）显示：(1) 配位键的性质强烈地依赖于铁的 $d_{x^2-y^2}$ 轨道的形式占据数。当这个数字为 0 时，$d_{x^2-y^2}$ 对 Fe-N 键的共价化有显著影响。其中，$d_{x^2-y^2}$ 轨道表现出一种与配位键的性质有关的有趣的行为，同时还表现出重叠集居对于轨道形式占据数的依赖关系。当 $d_{x^2-y^2}$ 的形式占据数为 0 时，$d_{x^2-y^2}$ 和N轨道间的重迭集居大约是 0.07。(2) 对于中性态，在铁原子上净电荷的范围在 1.39 至 1.53 之间（见表 10-16）。(3) Fe 的净电荷也取决于 $d_{x^2-y^2}$ 轨道的形式占据。

表 10-15 SCF 总能量和d电子组态

状态		形式占据数				SCF 总能量(a. u)
		d_{xy}	d_π	d_{z^2}	$d_{x^2-y^2}$	
低自旋居间自旋	$^1A_{1g}$	2	4	0	0	−2243.6697
	^3E_g(B)	1	3	2	0	.6691
	$^3B_{2g}$	1	4	1	0	.6991
	^3E_g	2	3	1	0	.7215
	$^3A_{2g}$	2	2	2	0	.7351
高自旋	$^5B_{1g}$	1	2	1	2	.6822
	$^5B_{2g}$	2	2	1	1	.7681
	^5E_g	1	3	1	1	.7694
	$^5A_{1g}$	1	2	2	1	.7802
π-阳离子	$^4E_u1$		$(^3E_g-1a_{1u})^+$			.5017
	$^4E_u2$		$(^3E_g-5a_{2u})^+$			.4851
	$^4A_{2u}$		$(^3A_{2g}-1a_{1u})^+$			.5151
	$^4A_{1u}$		$(^3A_{2g}-5a_{2u})^+$			.4979
d-阳离子	$^4A_{2g}$		$(^3E_g-2e_g)^+$			.5305

2. 基态和分子轨道

迄今为止，有关非轴向配位的铁-卟啉基态电子组态仍是有争议的问题（读者可参阅文献[63—65]）。其中，一种可能的基态电子组态是 $^3A_{2g}(d_{xy})^2(d_\pi)^2(d_{z^2})^2$，而另一种可能的基态电子

表 10-16　总原子轨道集居和 Fe-N 重迭集居

原子轨道	$^1A_{1g}$	^3E_g	$^3A_{2g}$	$^5B_{2g}$	^5E_g	$^5A_{1g}$
总原子轨道集居						
Fe　s	6.19	6.25	6.35	6.25	6.24	6.26
p_σ	8.07	8.05	8.04	8.07	8.05	8.03
p_π	4.09	4.09	4.09	4.11	4.10	4.10
d_{xy}	1.97	1.97	1.97	1.96	0.99	1.00
d_π	3.97	2.99	2.01	2.01	2.99	2.01
d_{z^2}	0.05	0.99	1.90	1.00	0.99	1.90
$d_{x^2-y^2}$	0.23	0.24	0.25	1.10	1.11	1.12
净电荷	1.44	1.42	1.39	1.50	1.53	1.49
N　σ	6.12	6.11	6.10	6.13	6.14	6.14
π	1.48	1.49	1.50	1.49	1.49	1.49
净电荷	−0.60	−0.60	−0.60	−0.62	−0.63	−0.63
Fe-N 重迭集居						
Fe　s	0.048	0.050	0.036	0.058	0.053	0.035
p_σ	−0.047	−0.047	−0.049	−0.038	−0.046	−0.047
p_π	0.043	0.041	0.040	0.040	0.044	0.042
d_{xy}	0.002	0.001	0.000	0.004	−0.001	−0.001
d_π	−0.001	0.002	0.004	0.004	0.001	0.004
d_{z^2}	0.013	0.012	0.020	0.009	0.011	0.021
$d_{x^2-y^2}$	0.066	0.067	0.060	−0.006	−0.003	−0.002
总和	0.124	0.127	0.121	0.076	0.055	0.053

表 10-17　^3E_g态分子轨道的组成和轨道能量(轨道能量的范围：−0.53至−0.16 a.u)

分子轨道	占据数	组分 %					
		Fe-d	Fe-s,p	N	C	H	轨道能
$6a_{2g}$	2	0	0	4	76	20	−0.525
$15a_{1g}$	1	95	5	0	0	0	−0.516
$2e_g$	3	98	0	0	2	0	−0.479
$19e_u$	4	0	3	55	41	1	−0.475
$9b_{1g}$	2	8	0	46	46	0	−0.473
$9b_{2g}$	2	82	0	4	10	4	−0.464
$1b_{1u}$	2	0	0	0	100	0	−0.449
$3e_g$	4	0	0	3	97	0	−0.383
$4a_{2u}$	2	0	2	31	67	0	−0.372
$4e_g$	4	1	0	38	61	0	−0.365
$2b_{2u}$	2	0	0	36	64	0	−0.355
$5a_{2u}$	2	0	2	20	78	0	−0.253
$1a_{1u}$	2	0	0	0	100	0	−0.235
$5e_g$	0	1	0	15	84	0	−0.162

组态是 $^3E_g(d_{xy})^2(d_\pi)^3(d_{z^2})^1$ 等等。Kashiwagi 等[58]根据对 FeTPP*的穆斯堡尔谱所作的理论分析认为,Fe-卟吩的基态是 ^3E_g。

表 10-17 和表 10-18 分别给出计算得到的^3E_g 态分子轨道的组成和轨道能量以及 ^3E_g 态的内壳层轨道能量。

从表中可以看出:(1) 最高占据轨道是 1 a_{1u},它是单纯的卟吩 π 分子轨道。(2) 1 a_{1u} 和 5 a_{2u} 轨道能量分别是 −0.235 和 −0.253 a.u,1 a_{1u}和 5 a_{2u} 的能量间隔是 0.49 eV,两者都是卟吩的 π 轨道。(3) 15 a_{1g}、2 e_g 和 9 b_{2g} 等三个分子轨道主要由铁原子的 3 d 轨道组成,它们分别占 95%、98%和 82%。

表 10-18 ^3E_g 态的内壳层轨道能量(a.u)

主要成分		轨道能
Fe	1 s	−261.353
	2 s	−31.656
	2 p_σ	−27.352
	2 p_π	−27.347
	3 s	−4.039
	3 p_σ	−2.649
	3 p_π	−2.646
N	1 s	−15.580
*$C_{\alpha\text{-pyr}}$	1 s	−11.296
$C_{\beta\text{-pyr}}$	1 s	−11.255～.253
Cmeso	1 s	−11.268

* 吡咯环的 α 碳原子和吡咯环的 β 碳原子

3. 电离

表 10-19 给出用ΔSCF 法计算的 Fe-卟吩^3E_g 的电离势。表 10-19 则给出了伴随着电离而出现的电子分布的变化。

从计算的重排作用能 (reorganization energy) 看,当任何一个电子从 Fed_π 轨道移开的时

表 10-19 Fe-卟吩 ^3E_g的电离势(eV)

轨道	阳离子态	重排作用能	ΔSCF 计算	观测的	
				FeOEP*	FeTPP**
1 a_{1u}	4 E_u1	0.41	5.98	6.06	6.50
5 a_{2u}	4 E_u2	0.43	6.43	6.48	6.80
2 $e_g(d_\pi)$	4 A_{2g}	6.15	5.20		

* 见文献[66]

** 见文献[67]

* FeTPP——铁-四苯基-卟吩。

候,此值是很大的。但是,当电子从卟吩的 π 轨道移开的时候,此值是小的。

从表中可以看出,在卟吩的 π 电子电离的情况下,变化几乎被限制在卟吩环内。另一方面,在铁的氧化中有相当数量的电子电荷从卟吩移动到铁上,因此,铁上电子电荷的净减少仅只是0.33。由此可见,在金属的氧化作用中,卟吩大环起着电子"缓冲器"的作用。

表 10-20 电离时电子分布的变化

原子		^3E_g	$^4E_u1-^3E_g-1a_{1u}$	$^4E_u2-^3E_g-5a_{2u}$	$^4A_{2g}-^3E_g-2e_g(d_\pi)$
Fe		24.58	−0.04	0.04	−0.33
N	×4	30.41	0.13	−0.10	0.29
$C_{\alpha-pyr}$	×8	47.90	−0.53	0.24	−0.19
$C_{\beta-pyr}$	×8	48.85	−0.14	−0.14	−0.14
C_{MESO}	×4	24.73	0.11	−0.52	−0.12
*H_{pyr}	×8	6.46	−0.36	−0.29	−0.34
H_{MESO}	×4	3.07	−0.14	−0.16	−0.14

* 吡咯环的氢原子。

除了 Almlöf 等的理论计算外,值得提及的还有 Rawlings 等的工作,他们对卟啉配合物的电子态作了较系统的研究[61]。Rawlings 等在用 ab initio 法计算铁卟啉配合物的低能电子态以及用 ab initio 组态相互作用计算$Fe^{II}(P)(NH_3)_2$的单线态、三重态和最低五重态的基础上,将理论计算结果跟四配位 Fe^{II}(TPP)、五配位高自旋和六配位低自旋正铁和亚铁血红素蛋白质的吸收光谱作了比较,获得了较满意的结果。Sontum 等[58]还用 X_α 多重散射法对氧合血红蛋白模型作了理论计算,并就计算结果对 Fe-配位键的性质和电子光谱进行了讨论。所有这些对金属-卟啉配合物电子结构与功能关系所作的研究,在不同程度上为我们提供了一些有意义的信息。如果将 Pullman 的早期工作及后来的 ab initio 等工作加以综合,则不难看出,通过卟啉配合物分子轨道的计算(事实上这样的计算还只能说是初步的),已经使我们能够在电子水平上对铁-卟啉配合物的某些生物学功能有了轮廓性的认识(见下一节)。

10.6 金属离子生物学功能的一些量子生物学观点

在前面,我们已经讨论过核酸、蛋白质和酶中金属离子的一些生物学功能。其实,金属离子在生物体内既能对一系列反应执行控制,同时它们又被许多其它反应所控制,从而产生错综复杂但又高度有序的反应体系。在这当中,金属离子与生物大分子间的相互作用又占据着十分重要的地位。问题非常明显,即我们不可能对这些反应体系的所有方面进行量子力学计算,尤其是在目前,或许我们只能对其中的某些局部问题作出量子生物学解释。

10.6.1 金属离子与生物大分子相互作用的若干问题

10.6.1.1 半导体掺杂观点

Szent-Györgyi 学派曾假设过两种最重要的生物大分子——蛋白质和核酸,是半导体[66-74]。

这一思想自 40 年代提出后，Szent-Györgyi 学派就不断地设法从禁带宽度计算方面来阐明这一思想。如果说，通过理论计算和实验，对蛋白质或核酸的能带宽度可以得到类似于本征半导体的结果[75]。那末，微量金属离子就可以认为是在这种蛋白质或核酸本征半导体中的掺杂(doping)。我们知道，掺杂是微量的，但对于本征半导体的导电却起着决定作用。从这个意义上来看，金属离子虽然处于痕量数量级，但对于发挥蛋白质(酶)或核酸的功能都是必不可少的。因此，金属离子和生物大分子间的相互作用在电子层次上的研究途径之一，就可能是通过借鉴固体物理学的思想和方法来探讨微量金属元素对蛋白质或核酸的禁带宽度的影响及其在这些半导体中的作用。

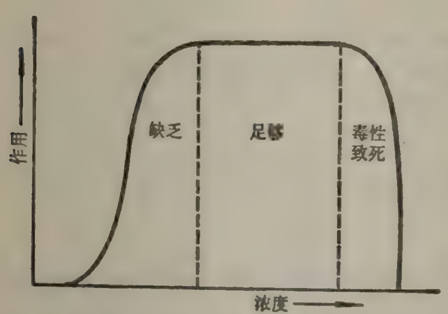


图 10-27 生物微量元素的浓度和生理作用

我们已经知道，在生命系统中，在一定浓度条件下，某些微量金属元素乃是生命活动所必须的，而一些元素则是有害的。Bertrand[76]早就指出，生命系统中的微量元素，在一定浓度范围内是有益的，而超过一定范围将是有害的(如图 10-27 所示)。对于这个道理，至今我们还仅仅是根据元素周期表及元素的核外电子壳层组态总结出一些唯象的规律。对于为什么某些元素在一定浓度下是生命活动所必需的，而在另一浓度下是有害的道理，还一无所知，更不必说在电子层次上的解释了。

根据上述的“掺杂”观点，我们是否可以把这些“必需”或“有害”的概念与它们对蛋白质或核酸禁带宽度的影响，对蛋白质或核酸本征半导体的导电机制的影响相联系起来呢？如果是这样，那也是很有意思的。至少：(1) 如果能够找到微量金属元素的生理作用与其对生物大分子的半导体导电性的关系或规律，则我们就可以从理论上去估计这些元素“有益”或“有害”的浓度范围。(2) 上述这种关系或规律的建立，将使我们有可能了解更多的微量金属元素的生理作用。

但是，目前的问题是，蛋白质或核酸半导体性的计算(参见 10.4.1)，还未得到能够说明问题的确切结果。虽然近年来，采用了更好的基集(双 ζ + 极化函数)、应用了改进的电子极子模型[77]，在计算方法上作了很多改进，使计算的禁带宽度从 8 eV 降至 3 eV 左右(肖奕和陈润生等计算的 poly(Gly-Na$^+$)I，poly(Gly-Na$^+$)II 的禁带宽度分别为 10.421 eV 和 9.524 eV)，但本征半导体性的问题仍未解决[75]。因此，首要的问题是，对于诸如蛋白质和核酸这样的生物大分子的本征半导体性的计算要得到满意的结果，然后才可能考虑金属离子对其所产生的影响。近来，运用涉及到非线性微分方程的孤波子(Soliton)理论对于低维周期性聚合物分子的固体物理学计算有了很大进展，因此，也期待着这种方法进一步渗透到生物大分子的理论计算上来而有所突破。

10.6.1.2 电子-构象相互作用观点

与上述 Szent-Györgyi 学派不同，Volkenstein[78]持有另一种观点。他认为，由于计算的蛋白质禁带宽度在 5.0 eV 以上，故蛋白质不是半导体。再者核酸的糖-磷酸酯骨架和蛋白质的肽

链都不是共轭体系。因此,生物聚合物分子的生物学作用不可能由它们的半导性或电子在整个分子体系中的离域作用决定。Volkenstein、Vallee 等[79-81]还认为,蛋白质(酶)的功能在于其高级结构,即在于其构象。所以,他们提倡应从电子-构象相互作用 (electron-conformation interaction,ECI)上去探求生物聚合物分子的活性。已知有约1/3的酶是金属酶或金属离子激活酶,且金属离子对酶的构象变化有直接作用,ECI是酶作用的基础。从前面的讨论中,我们已经知道。金属离子与配位体(如蛋白质分子)相互作用对金属离子的催化特性有十分明显的影响。

一般地,在金属酶中金属离子的作用为:

(1) 作为 Lewis 酸,电子对供体;

(2) 氧化-还原过程的催化剂,电子传递体;

(3) 支持酶的结构。

现在,我们来分析一个作为 Lewis 酸的金属离子催化的反应和金属酶催化的反应的例子。图 10-28 是二价金属离子催化的丙酮二羧酸酯的分解反应的速度常数和丙二酸二价金属配合物稳定常数的相关关系。其催化能力的大小和多数模型反应的催化强度次序一致:

$$Ca<Mg<Mn<Fe<Co$$
$$\leqslant Zn<Ni<Cu$$

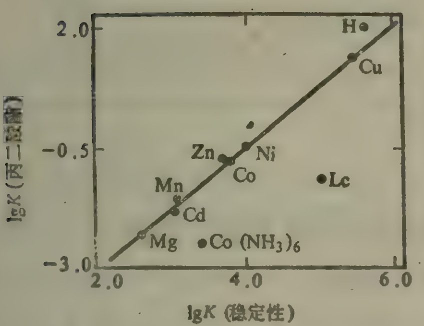


图 10-28 以丙二酸为模型的二价金属离子的催化反应速度常数和配合物稳定常数的相关关系

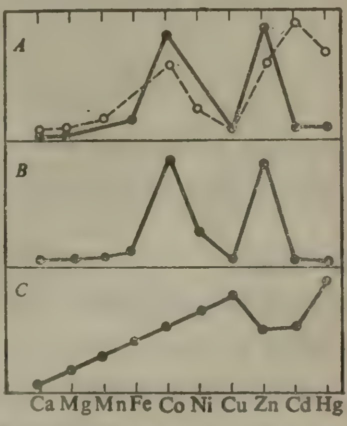


图 10-29 金属酶催化的反应之相对速度
A. ——羧肽酶作肽酶 ------羧肽酶作酯酶
B. 碳酸酐酶
C. 模型体系

而这一次序和稳定常数次序一致,且显示了很好的线性自由能关系。稳定常数是金属-配位体键的强度的量度。显然,这些二价金属离子的催化活性及其与配位体的相互作用而成的键的强度直接相关。

图 10-29 表示二价金属离子的金属酶催化的反应之相对速度比较。

从图中可知,金属酶催化的催化强度次序和上述一般催化强度次序不一致。表 10-21 和表 10-22 列出了金属羧肽酶和金属碳酸酐酶的生成常数。图 10-30 给出了二价金属的 Irving-

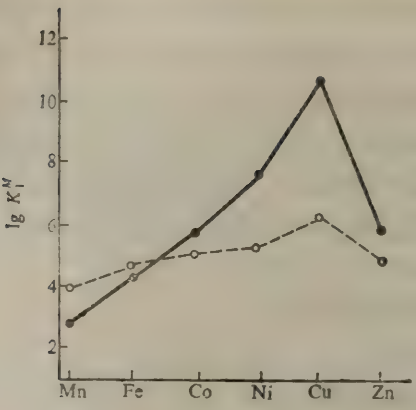


图 10-30 二价金属的 Irving-Williams 顺序
○-草酸盐，●-1，2-二氨基乙烷

表 10-21 金属羧肽酶的生成常数

金属 Metal	LogK 表观 pH 8.	LogK 校正 pH 8.0
MnII	5.6	5.6
CoII	5.8	7.0
NiII	5.7	8.2
CuII	5.1	10.6
ZnII	8.3	10.5
CdII	7.9	10.8
HgII	6.7	21.0

表 10-22 金属碳酸酐酶的生成常数

金属	LogK，pH5.5
MnII	3.8
CoII	7.2
NiII	9.5
CuII	11.6
ZnII	10.5
CdII	9.2
HgII	21.5

Williams 顺序。

通过分析比较可以看出，(1) Co^{2+} 和 Zn^{2+} 的催化活性在金属羧肽酶(图 10－29 *A*)和金属碳酸酐酶(图 10－29 *B*，参见 10.4.2)中最高，这与模型体系(图 10－29 *C*)反应的催化强度不同，但后者的催化强度次序却与 Irving-Williams 顺序一致(见图 10－30)。(2) 在金属羧肽酶和金属碳酸酐酶中，金属离子和酶蛋白(apoenzyma)的结合稳定常数也与其相应的金属酶催化强度次序不一。这些变化是不是由于不同的金属离子和蛋白质(酶)上的不同的官能团相结合而引起的？例如

$$\overbrace{Ca,Mg,Mn}^{-O^-}\quad \overbrace{Fe,Co,Ni}^{\equiv N,-O^-}\quad \overbrace{Cu,Zn}^{\equiv N,S^-}\quad \overbrace{Cd,Pb,Hg}^{-S^-}$$

然而研究表明，金属酶上的金属离子所处的活性部位是精确地固定的，这一位置并不因其它金属离子取代了原来的金属离子而改变。因此，排除了不同金属离子和不同官能团相结合的那种可能性。

由于酶和底物的结合是一个涉及构象因素的中间过渡态，因此，Volkenstein 等认为，金属酶中金属离子催化活性的不同是由于金属离子和构象的相互作用，形成了一种不同于基态构象的

过渡态构象。金属离子改变了酶的构象，反过来酶的构象状态也决定了金属离子的电子状态，且不同的金属离子本身又有不同的独特的电子组态。因此，不同的金属离子与酶构象的相互作用是不同的。一种解释认为，Zn^{2+}、Co^{2+}的较高催化活性是由于两者较易使酶的基态构象变形而形成一种有利于反应的中间态构象所致。从目前的发展趋势看，ECI 理论很可能成为研究金属离子与蛋白质或核酸相互作用的有力工具。为此，以下我们将简要地介绍一下 Velkenstein学派对 ECI 理论的研究[79,82-86]。

Velkenstein 学派认为，酶不仅起催化作用，而且还是一种特殊的反应介质。低介电常数的介质可使反应物之间有较强的电相互作用。酶内部的非极性区域为细胞提供了类似于有机溶剂的低介电常数介质的反应场所，因此底物在酶的非极性介质环境里可以与酶的极性基团之间有较强的电相互作用。在酶-底物复合物里，反应物处于一种非极性的或部分极性的介质之中，以酶环境为介质对反应物的作用和以溶剂为介质对反应物的作用其物理本质是一样的，因此可以按电介质或非电介质溶剂的反应动力学理论来构建一个把酶内环境视作反应介质的反应理论。在谐振近似下，介质的 Hamilton 量如下：

$$H=-\frac{1}{2}\sum_{x}\hbar\omega_{x}\frac{\partial^{2}}{\partial q^{2}_{x}}+\frac{1}{2}\sum_{x}\hbar\omega_{x}(q_{x}-q_{xo})^{2}+I \tag{10.11}$$

式中，ω_x 和 x_o 分别是介质正常振动的频率和坐标，I 是系统之最小势能(或自由能)。一般而言，介质的平衡组态必须用活性部位的不同的构象状态 α 和不同的电子状态 n 来表征。

当考虑电子-构象相互作用时，介质的 H 应写成：

$$H=-\frac{1}{2}\sum_{x}\hbar\omega_{x}^{\alpha,n}\frac{\partial^{2}}{\partial q^{2}_{x}}+\frac{1}{2}\sum_{x}\hbar\omega^{\alpha,n}(q_{x}-q_{xo}^{\alpha,n})^{2}+I \tag{10.12}$$

而反应的一个基元作用的机率是：

$$w=\sum_{\alpha\alpha,nn'}\phi_{\alpha n}w_{\alpha n,\alpha' n'} \tag{10.13}$$

这里 $\phi_{\alpha n}$ 是介质在反应前其构象状态为 α 及其电子态为 n 的机率，$w_{\alpha n,\alpha' n'}$ 是其过渡到状态 $\alpha' n'$ 时的机率。对应于不同构象和电子状态的位能面如图 10－31 和图 10－32 所示。

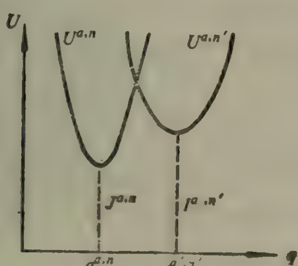


图 10－31　构象态 α 无变化时的电子迁移

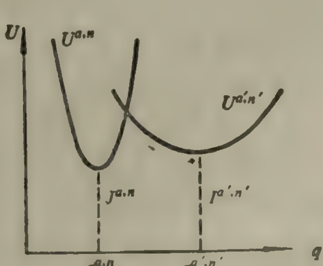


图 10-32　构象态 α 有变化时的电子迁移

极性介质和非极性介质之间的差别是由于电子的转移或电子密度的变化可以伴随一较大的熵变。表征极化涨落的频率在反应过程中並不改变，但是表征变形的声频却可有较大的变化。因此，在非极性介质中的反应速率可受活化熵的影响。由于在非极性介质中声频的值较低，所以

电子体系总是绝热地跟着变形振动而变化，且透射系数 K 必定等于1。

ECI 的一般意义是酶反应过程中的大分子的构象转变出现在电子组态的重组之后，这种情况可用固体物理学的概念来描述。我们知道，金属离子是许多酶的辅因子，因此可以认为，利用金属酶来研究 ECI 是十分合适的，反过来讲，ECI 又是研究金属酶作用的一种方法。

10.6.2 金属离子与氧的积累、贮存和输运

过渡金属配合物与氧分子的配位结合不仅是氧的活化的基础，而且也是积累、贮存和输运氧的基础。研究已经证实，完成贮存和输运氧的功能的最重要的蛋白质是含铁蛋白质（如肌红蛋白和血红蛋白等）、含铜蛋白质和含钒蛋白质。

Яцимирский 认为[2]，要说明氧分子与这些生物配合物相互作用的机理，必须探讨 O_2^{z-} 型粒子的重要特征（表 10-23）。

表 10-23 O_2^{z-} 型粒子的特征

粒子	占据轨道			谱项及它们的能量(eV)	O—O间的距离(Å)	亲核能(kcal/mol)
	$\sigma_g(p_z)$	$\pi_u(p_xp_y)$	$\pi_g(p_xp_y)$			
氧型O_2	2	4	2	$^3\Sigma_g{}'\Delta_g(23.4)$ $^1\Sigma_g(37.5)$	1.21	119
超氧型O_2^-	2	4	3	$^2\Pi_g$	1.28	351
过氧型O_2^{2-}	2	4	4	$^1\Sigma_g$	1.49	529

氧分子基态的两个 π^* 轨道上各有一个电子，所以是三重态——$^3\Sigma_g$。

金属离子与氧分子的可逆反应有两条途径：

$$M^{n+} + O_2 \rightleftharpoons M^{n+}O_2$$

或

$$2\,M^{n+} + O_2 \rightleftharpoons M_2{}^{n+}O_2$$

为了使反应能够进行，必须确保一定的而且是很严格的条件，即必须满足金属轨道在能量和对称性方面的要求。这是因为氧分子具有很复杂的结构，它有三个不同对称性的轨道：σ_g、π_u 和 π_g。其中 σ_g 和 π_u 轨道为给电子轨道，π_g 轨道为接受电子的轨道。由此可见，O_2 在形成配合物时，既可作为电子给予体，也可作为电子接受体。因此，只有当金属离子的轨道具有这样的两重性（π 电子给予体、π 电子接受体）时，才能与 O_2 形成稳定的配合物。由 10.5.1 和 10.5.2 的讨论知，离子型亚铁和正铁卟啉既是较好的 π 电子给予体又是较好的 π 电子接受体。

O_2 与五配位亚铁配合物 FeL_5 配位时，Fe(II)的 d_{z^2} 轨道和 d_{xz} 轨道分别与 O_2 的 σ_g (π_u)和 π_g 轨道相互作用（图 10-33），由于 π_g 轨道有两个电子空位，$Fe \rightarrow O_2$ 反馈使氧配合物稳定化。

由于反馈 π 键决定 O_2 结合能力，所以只有低氧化态过渡金属离子才有可能与 O_2 配位。不过，低价态金属离子又容易被 O_2 氧化，于是，氧合与氧化成为一对矛盾。作为氧载体，要求配合

物只氧合而不氧化，反之作为电子传递体，要求氧化而不氧合。

对于 O_2 的配位研究过两种情况，一是 O—O 的分子轴位于卟吩的 x 轴共平面的二个 N 原子之上；另一种情况是 O—O 的分子轴位于与 z 轴共轴的铁原子之上。d_{xy} 由于与 O_2 的 π 轨道相结合而极大地稳定，在这种状态下铁的电荷为 +0.30，接近 Fe(III)的状态。与此相对应，O_2的电荷为 −0.53。O_2 在共轴配位的情况时，它的 π轨道能有很大下降，并且由于这种未充满轨道，使电子极易在其中的 d_π 轨道移动，



图 10-33

从而呈现不稳定状态(例如血红素，通过共轴配位通常影响到金属 d 轨道能量)。因此推测 O_2 是共平面配位。

在人和大多数动物中，氧载体为血红蛋白(Hb)，贮氧体为肌红蛋白(Mb)。Hb 与 Mb 的基本结构相同，都是由一个血红素单元与一个珠蛋白链结合成的分子构成的。所不同的是 Mb 只含一个血红素蛋白链，而 Hb 则是由四个亚单位组成的四聚体，每一个亚单位都类似于 Mb 但又不等于 Mb。因而 Hb 具有血红素-血红素间的相互作用造成的协同作用，而 Mb 缺乏这种性质。

血红蛋白分子既能够与 O_2 迅速结合，也能够与 O_2 迅速解离。其结合与解离主要决定于血液中 O_2 分压的高低。O_2 在血红蛋白分子中的结合部位是血红素中的 Fe^{2+}。在氧合血红蛋白(HbO_2)中，血红素的 Fe^{2+}，其配位数是六，其中每个亚单位结合一个 O_2，故每个血红蛋白分子共结合四个 O_2。血红素铁是分子呼吸的“触发器”。没有与 O_2 结合的血红蛋白(Hb)中，铁的配位数为 5，其中四个来自卟啉环的 N，另一个来自近侧组氨酸的 3 位 N。此时配位场较弱，铁处于高自旋态，故其离子半径较大。突出在血红素卟啉环平面之外。与此对应，氧合 Hb，血红素的 Fe^{2+}，其第六个就是O_2 的结合位置。分子氧使总配位场的强度增大，此时 Fe 处于低自旋态，其离子半径比脱氧时缩小 13 %，于是“落入”卟啉环平面之内(参见图 10-34)。这一空间变化通过

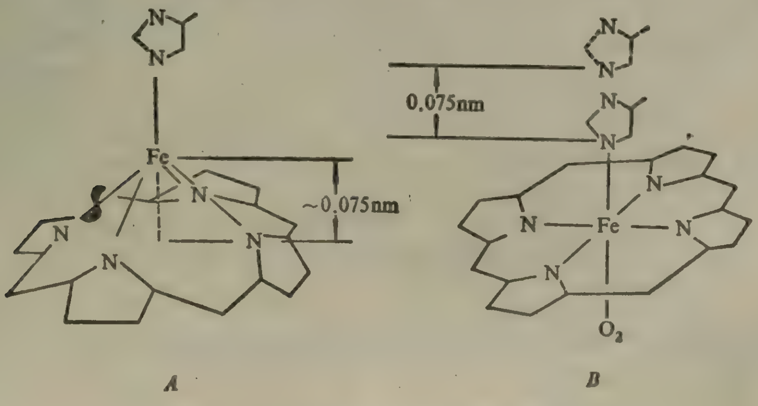


图 10-34 去氧和氧合血红蛋白血红素部分的起落

与之配合的组氨酸残基牵动与之相联的蛋白质部分发生显著移动，并作为分子呼吸的"触发器"，引起整个血红蛋白分子发生构象重排。

对 Hb 和 Mb 的载氧贮氧功能与结构的关系已经做过大量研究，其中包括对这两种生物分子的直接研究和对模型化合物的研究。

Otsuka 等[87]曾计算过 Hb 和 Mb中血红素铁原子的电子结构(参见图 10-35)。

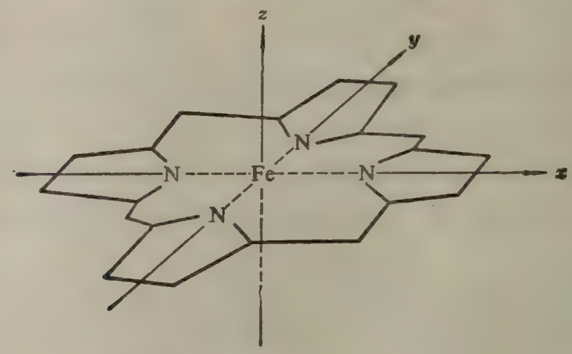


图 10-35

如果将铁定为原点，那么铁原子的 5 个d 轨道就可表示为：

$$u=\frac{1}{2\sqrt{3}}R(r)\frac{3z^2-r^2}{r^2},\quad v=\frac{1}{2}R(r)\frac{x^2-y^2}{r^2},$$

$$\xi=R(r)\frac{yz}{r^2},\quad \eta=R(r)\frac{xz}{r^2},\quad \zeta=R(r)\frac{xy}{r^2} \tag{10.14}$$

在亚铁和正铁肌红蛋白中，铁的 d 电子的轨道能如图 10-36 A 和 10-36 B所示。

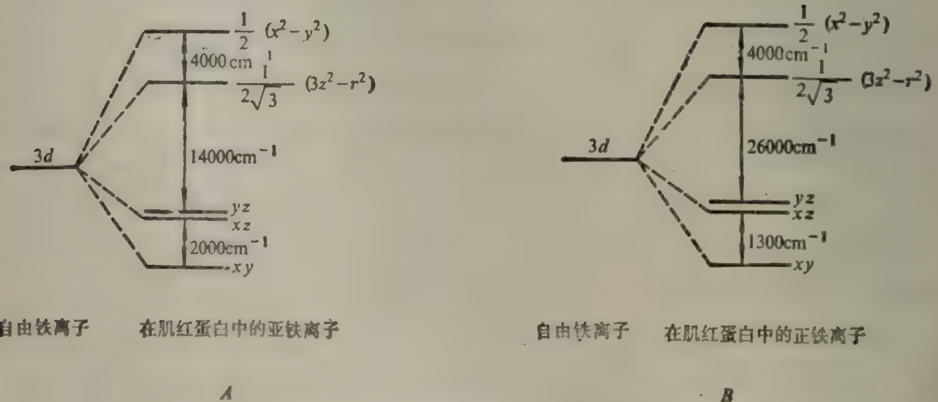
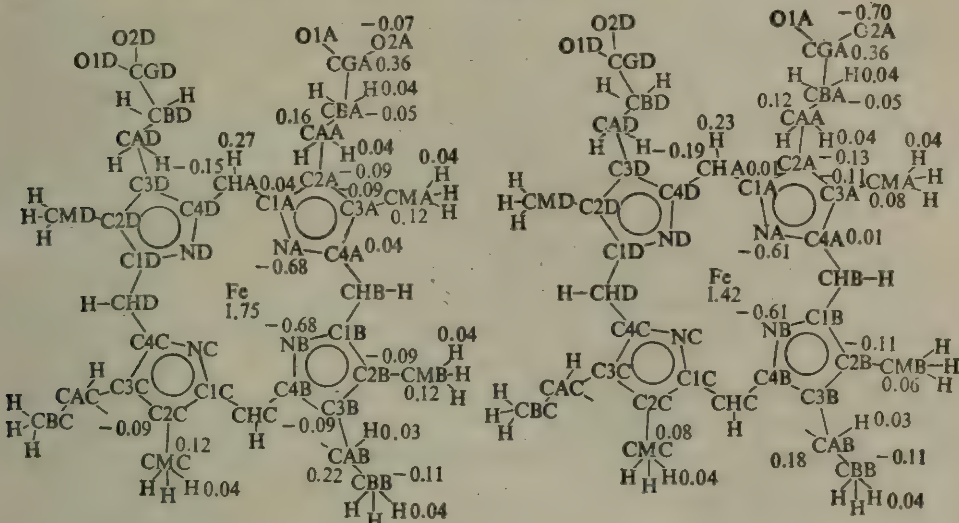


图 10-36 在亚铁和正铁肌红蛋白中，铁的 d 电子轨道能

已经知道，血红蛋白、肌红蛋白和细胞色素等等都含有血红素基，且后者又是生物化学反应中的关键结构。因此，历来受到生物学家的重视。但是，对于量子生物学家(或量子化学家)，要处理这样的大分子在实际上是非常困难的。实践已经证明，用简单的方法是不可能克服理论计

算中的困难的。近年来，发展了处理血红素基电子结构的一些方法，其中 Pèpe 等[88] 计算了氧化和还原型的血红素基的净原子电荷(图 10-37)，获得了较好的效果，不仅如此，Pèpe 等还在分子静电势计算的基础上，着重讨论了细胞色素的活性。通过对血红素基电子结构的理论计算看到，在血红素的氧化型中，静电势在 Van der Waals 等值线处处是正的。另一方面，在还原型中，分子的非血红素部分对血红素分子静电势的影响相当大。对这些计算结果所作的分析表明，对细胞色素而言，非血红素部分在电子传递反应中起着催化作用。



氧化型血红素基　　　　还原型血红素基

图 10-37　氧化和还原型血红素基的净原子电荷

我们知道，细胞色素是一类含血红素的酶类，参与氧化还原反应，在线粒体的电子传递链中起着重要作用。在此链的末端部分有五种具有不同氧化还原电位的细胞色素，按氧化还原电位增加的顺序是：细胞色素 b、细胞色素 c_1、细胞色素 c、细胞色素 a 和细胞色素 a_3。它们依次地接受和传出电子，最后被分子氧再氧化，并用环境中的质子形成水，同时将氧化还原反应中产生的能量储存在 ATP 中。图 10-38 表示随着 ATP 生成的电子传递系统的一般模式图，其中 X、Y、Z 表示相关的酶。

迄今，尽管对此电子转移的机制进行了一系列的研究，但仍不十分清楚。

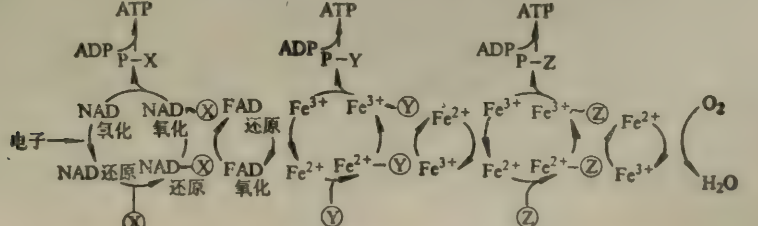


图 10-38　随着 ATP 生成的电子传递系统的一般模式图

金属离子与氧的积累、贮存和输运的研究内容相当丰富。我们仅只就量子生物学的研究，作了简单的讨论。近几年来，这一领域的研究已经取得了明显的进展，可望今后在量子生物学研究中取得实质性成果。

参 考 文 献

[1] Чанс, В.: "Перенос электронов в биологических системах", электроника и кибернетика В биологии и медицине",(издательство иностранной литературы, москва), (1963).
[2] К. Б. 雅茨米尔斯基著，余珣、吴瑶曼译：«生物无机化学入门»，科学出版社(1985).
[3] Theophanides, T.: Inter. J. Quantum Chem., 26:933(1984).
[4] Pullman, A. et al.: Biochem. Rev., 50:118(1980).
[5] 江寿平、乐树云等：生物化学与生物物理学报，13:485(1981).
[6] Melius, C. F. et al.: J. Chem. Phys., 56:3342(1972).
[7] Dolphin, D. et al.: Inter. J. Quantum Chem., 16:311(1979).
[8] Pullman, A. et al,: Chem. Phys. Lett., 47:411(1977).
[9] Perahia, D. et al.: Theor. Chim. Acta, 43:207(1977).
[10] Pullman, A. et al.: Inter. J. Quantum Chem., 10:59(1976).
[11] Hehre, W. G. et al.: J. Chem. Phys., 52:2769(1970).
[12] Шусторович, Е. М.: "Химическая связы в координационных соединения," (знание, москва), (1975).
[13] М.И. 休斯著，陈自在等译：«生物无机化学»，农业出版社 (1986).
[14] 周公度著：«无机结构化学» (无机化学丛书)，科学出版社(1982).
[15] 张祥麟等编著：«配位化学»，中南工业大学出版社(1986).
[16] 王夔：分子科学与化学研究，2:133(1982).
[17] 唐敖庆等著：«配位场理论方法»，科学出版社(1979).
[18] Perahia, D. et al.: Theor. Chim. Acta, 43:207(1977).
[19] Kistenmacher, T. J. et al.: "Metal-Ligand Interactions in Organic Chemistry and Biochemistry", Pullman, B. (Ed) (D. Reidel Publishing Company, Dordrecht-Holland/Boston-U. S. A.), (1977).
[20] Coulter, C. L.: J. Am. Chem. Soc., 95:570(1973).
[21] 唐明生、刘若庄等：生物化学与生物物理学报，18:474(1986).
[22] Schildkraut. C. et al.: Biopolymers, 3:195(1965).
[23] Deve, W. F. et al.: J. Mol. Biol., 5:467(1963).
[24] Karpel, R. L. et al.: Biochem., 19:504(1980).
[25] Pohl, F. et al.: J. Mol. Biol., 67:375(1972).
[26] Haulon, S. et al.: Biochem. Biophys. Acta, 519:526(1978).
[27] Chan, A. et al.: Biochem., 18:84(1979).
[28] Reuben, J. et al.: Proc. Natl. Acad. Sci. U. S. A., 71:245(1975).
[29] Pack, G. R.: Inter. J. Quantum Chem., Quantum Biol. Symp., 9:81(1982).
[30] Liebmann, P. et al.: J. Am. Chem. Soc., 104:691(1982).
[31] Noyes, R. M.: J. Am. Chem. Soc., 84:513(1962).
[32] Karenzi, P. C. et al.: Biopolymers, 9:181(1979).
[33] 肖奕、陈润生等：生物物理学报，2:51(1986).
[34] 肖奕、陈润生等：科学通报，30:1418(1985).
[35] Rozsnyai, B. F. et al: J. Chem. Phys., 53:4325(1970).
[36] Kannan, K. K. et al.: Cold Spring Harbor Symp., Quantum Biol., 36:221(1971).
[37] Kannan, K. K. et al.: Proc. Natl. Acad. Sci. U.S.A., 72:5(1975).
[38] Jonsson, B. et al.: J. Am. Chem. Soc., 99:4628(1977).
[39] Jonsson, B. et al.: J. Am. Chem. Soc., 100:1658(1978).

[40] Magid, E. et al.: Biochim. Biophys. Acta, 165:515(1968).
[41] Pinset, B. R. et al.: Trans. Faraday Soc., 52:1512(1956).
[42] Coleman, J. E.: "Inorganic Biochemistry", Eichhorn, G. L. (Ed.) (Elsevier, Amsterdam), (1975).
[43] Sawaryn, A. et al.: Inter. J. Quantum Chem.,16:293(1979).
[44] Pullman, A. et al.: Inter. J. Quantum Chem.,16:641(1979).
[45] Kenner, G. W.: 生物科学动态，(10):23(1963).
[46] Granick, S. et al.: Adv. Enzym. VII, Interscience Publishers, Inc., New York, (1947).
[47] Pullman, B. et al.: "Quantum Biochemistry", (Wiley, Interscience, New York, London), (1963).
[48] 永田亲义著，陶宗晋、江寿平译：«量子生物学入门»，上海科学技术出版社(1979).
[49] Spanjaard, C. et al.: J. Chim. Phys.,:169(1961).
[50] Pullman, B. et al.: Proc. Natl. Acad. Sci. U.S.A., 46:1011(1960).
[51] Spanjaard, C.: "Diplome d'Etudes Supérieures", (Faculte des Science de Paris), (1960).
[52] Almlöf, J.: Inter. J. Quantum Chem., 8:915(1974).
[53] Kashiwagi, H. et al.: Inter. J. Quantum Chem., 20:843(1981).
[54] Dedieu, A. et al.: Nouv. J. Chim., 3:653(1979).
[55] Dedieu, A. et al.: Adv. Quantum Chem., 16:43(1982).
[56] Obara, S. et al.: J. Chem. Phys., 77:3155(1982).
[57] Nozawa, T. et al.: Bull. Chem. Soc. Jpn., 56:1721(1983).
[58] Sontum, S. et al.: J. Chem. Phys., 79:2881(1983).
[59] Zemer, M. et al.: Theor. Chim. Acta, 4:44(1966).
[60] Loew, G. et al.: "Iron Porphyrins", Lever, A. (Ed.) (Addison-Wesley, Reaging, Mass.) Part I, (1983).
[61] Rawlings, D. et al.: Inter. J. Quantum Chem., 28:773(1985), 28:797(1985), 28:823(1985).
[62] Dedieu, A. et al.: "Metal-Ligand Interactions in Organic Chemistry and Biochemistry", Part 2 (Reidel, Dordrecht/boston), (1977).
[63] Collman, J. P. et al.: J. Am. Chem. Soc., 97:2676(1975).
[64] Kobayashi, H. et al.: Bull. Chem. Soc. Jpn., 45:450(1972).
[65] Goff, H. et al.: J. Am. Chem. Soc., 99:3641(1977).
[66] Szent-Gyorgyi, A.: Science, 93:609(1941).
[67] Szent-Gyorgyi, A.: Nature, 148:157(1941).
[68] Szent-Gyorgyi, A.: "Bioenergetics", (Academic Press, New York), (1957).
[69] Szent-Gyorgyi, A.: "Introduction to a Submolecular Biology", (Academic Press, New York), (1960).
[70] Szent-Gyorgyi, A.: "Bioelectronics," (Academic Press, New York), (1968).
[71] Szent-Gyorgyi, A.: "The Living State", (Academic Press, New York), (1972).
[72] Szent-Gyorgyi, A.: "Electronic Biology and Cancer," (Dekker, New York), (1976).
[73] Evans, M. et al.: Biochim. Biophys. Acta, 3:188(1949).
[74] Ciba Foundation: "Submolecular Biology and Cancer", (Excerpta, Amsterdam) (1976).
[75] Ladik, J. et al.: Inter. J. Quantunm Chem., Quantum Biol. Symp., 10:73(1983).
[76] Schroeder, H. A.: "The Trace Elements and Man", The Devin-Adair Comp., Old Greewich, CT., (1973).
[77] Suhai, S.: Phys. Rev. B, March 15, (1983).
[78] Volkenstein, M.: "Molecular Biophysics", (Academic Press, New York), (1977).
[79] Ibid. p. 392.
[80] Vallee, B. et al.: Chem. in Britain, 4:397(1968).
[81] Vallee, B. et al.: Proc. Natl. Acad. Sci. U.S.A., 59:498(1968).
[82] Volkenstein, M.: Izv. Acad. Sci. USSR Ser. Biol., No. 6:805(1971).
[83] Volkenstein, M.: J. Theor. Biol., 34:193(1972).

[84] Ivanov, V. et al.: Eur. J. Biochem., 40:519(1973).
[85] Volkenstein, M.: Pure Appl. Chem., 36:9(1973).
[86] Ivanov. V. et al.: Mol. Biol., 8:433(1973).
[87] Otsuka, J. et al.: "Electronics Structure of Heme Iron Atoms in Myoglobin and Hemoglobin", International Union for Pure and Applied Biophysics, Symp. Papers. Pushchino. (1973).
[88] Pèpe, G. et al.: J. Theor. Biol., 115:571 (1985).
[89] 理科化学教材编审委员会(物质结构编审组)：《物质结构教学文集》,高等教育出版社， 197—204(1983).

刘次全
谭载友
刘征先

第 11 章　生物系统中能量和电子的转移

在 3.2.3 中，我们讨论了有关“电荷迁移相互作用”的一些问题。鉴于电子的转移(或迁移，见后)在生物学过程中具有重要的作用，因此在这一章里，我们将较详尽地叙述有关“生物系统中能量和电子的转移”在电子水平上研究的一些进展。

在生物系统中，起基本作用的粒子是电子和质子。电子和质子的转移不仅和生物系统中所发生的反应密切相关，而且与生物系统的能量转移密切相关。甚至可以认为，这样的相关性在生物系统中起到了举足轻重的作用，例如光合作用和生物氧化等这样一些生物学过程就是如此。如果从电子水平看，则诸如生物大分子本身的稳定性、反应活性以及储存、传递能量和信息等方面的性质，都与大分子内部原子的运动，特别是电子的运动密切相关。

从目前的情况看，在生物系统中，大分子的能量转移作为“生物能力学”研究的一个方面以及就这一问题所开展的其它一些研究，取得了一些进展[1-4]。有关电子和质子转移的研究，同样取得了进展。正如上面所述，在生物系统中，电子、质子的转移和能量的转移密切相关，因而许多研究成果都反映了这种相关性。但是迄今为止，仍有许多问题未能获得令人满意的解释和解决。

我们看到，以量子力学作为理解生命过程的方法论的量子生物学，已经在通过媒介物和隧道效应的电子、质子转移；电子、质子转移和能量转移的偶联以及生物大分子内和分子间在能量、电子、质子转移的高效率的原因等方面的研究领域中取得了一些进展[4,5]。

另一方面，在分子水平上解决某些生物学问题的各种方法中，利用固体理论方法的尝试近年来引起了人们的注意。这些尝试是以核酸和蛋白质分子周期结构的实验资料为依据的。当研究生物能力学的中心问题之一——说明在一个大分子内以及分子间能量、电子和质子转移的高效率的原因时，也经常利用固体理论方法。

研究实践提示我们，对处于水环境中的复杂生物系统，为了对其能量和电荷转移现象进行理论研究，不可避免地要把反映这些现象的主要性质予以简单化和模型化。在这方面 Давыдов 做了较深入的工作[6]。

11.1　生物大分子的能量转移

能量转移是包括许多种类和性质各不相同的能量传递形式，是一种微观过程。能量转移(energy transfer)与能量迁移(energy migration)有时并不严格区分，一般可以通用（为便于讨论，我们统称“转移”）。但严格地讲，所谓能量迁移是指能量的多次转移。

我们知道，生物系统是一个热力学上的开放系统，在这个系统内存在着物质、能量和信息的持续不断的流动。其中能量的流动只能朝一个方向不可逆地进行。由于生物体系的能量输入是来自于太阳供给的可见光，能量输出是表征为热的近红外光。因此不难看出，这个体系的能量流是一种光子由高能谱向低能谱的不可逆流动。正是这种光子流构成了所有生物系统内的能量背景。

11.1.1 能量的共振转移

共振转移是电子激发能的转移方式之一，是发生在量子水平上的分子间相互作用。在这一节里，我们主要讨论激发能的无辐射的共振转移。

共振转移是一种相同分子或不同分子之间的长距离的能量转移，是一种长程作用。电子激发能本身可以看成是一种振动，只要吸收能量的分子(受体)的振动电场是在激发分子(供体)的振动电场之中，而且供体的荧光光谱与受体的吸收光谱有重叠部分，激发分子的电磁振动就会引起吸收分子的同频共振，从而转移电子激发能。归纳起来，发生能量共振转移需要满足以下三个条件：

(1) 能量供体必须是发荧光的(在没有受体存在时)，荧光量子产额愈高，能量转移过程愈有效；

(2) 供体的荧光光谱必须与受体的吸收光谱有重叠，光谱重叠愈大，能量转移效率也愈高；

(3) 要发生有效的能量转移，供体与受体间的距离不能超过 5～10nm。

共振转移既然涉及到供体和受体分子，从物理理论上就可把它近似地看成是振动的两个偶极子的相互作用，因此有人把共振转移叫做偶极-偶极相互作用的转移就是这个原因。它们之间的作用实际上就是激发的供体对受体部位的电磁场扰动的结果。

早在 1948 年 Förster 就提出了适用于供体和受体间能量相互作用的能量转移机制。随后，他又推导出表示这种能量传输规律的公式[6,7]。有关激发能借共振机制转移的理论基础，曾沿经典的和量子力学的路线这样两个方向发展。在非常弱的偶联情况下，供体和受体间的相互作用比供体能级的振动分裂要小，两个途径对于 $\eta_{A\to B}$都产生相同的表达式，$\eta_{A\to B}$是表示每个单位时间里从供体到受体的一个能量量子的转移机率：

$$\eta_{A\to B}=\frac{9000(\ln 10)k^2\phi}{128\,\pi^5 n^4 N r^6 \tau}\int_0^\infty \frac{f(\bar{\nu})\varepsilon(\bar{\nu})}{\bar{\nu}^4}\mathrm{d}\bar{\nu} \tag{11.1}$$

式中，$k=\cos\theta_{AB}-3\cos\theta_A\cos\theta_B$。$\theta_{AB}$是供体与受体偶极间的角度，$\theta_A$ 与 θ_B 分别为供体和受体与其连线之间的角度，ϕ 是在没有受体分子存在时的供体的量子产额，n 是介质的折射指数，N 是 Avogadro 常数，r 是供体与受体间的距离，τ 是在没有受体分子存在时供体的寿命，$f(\bar{\nu})\mathrm{d}\bar{\nu}$ 是供体在波数 $\bar{\nu}$ 到波数 $\bar{\nu}+\mathrm{d}\bar{\nu}$ 范围内规范化的荧光强度。因此

$$\int_0^\infty f(\bar{\nu})\mathrm{d}\bar{\nu}=1$$

而$\varepsilon(\bar{\nu})$是受体在波数 $\bar{\nu}$ 时的吸收系数。

将(11.1)式的右边写成$(1/\tau)(R_0/r)^6$，R_0 除 τ 与 r 外，含有体系中所有参数，所以，R_0 可表示为

$$R_0^6=8.8\times10^{-25}\phi k^2 n^{-4} J \tag{11.2}$$

J 是式(11.1)中的光谱重叠的积分项。

共振转移是与供体自发衰变竞争，后者的速率常数是 $1/\tau$。所以能量转移效率T 应为

$$T = \frac{k_T}{\tau^{-1} + k_T} = \frac{R_0^6}{r^6 + R_0^6} \tag{11.3}$$

R_0 的定义是：当供体有一半能量转移给受体，另一半能量是以其它形式(包括发光在内)消耗了，这时供体与受体间的距离。简言之，R_0 就是转移效率为 50%时的供体和受体间的距离，又叫做临界距离。从(11.2)式可见，R_0决定于 ϕ、k、n 与 J 等四个参量。能量转移速率愈快，R_0 的数值就愈大。Förster 关于激发能借偶极-偶极相互作用的无辐射转移的著名特点，就是它与 $1/r^6$ 成比例。Stryer 等[8-10]的实验证明了它们对供体与受体间的距离有显著的依赖性。

产生共振转移的可能性(几率)还取决于电子激发态的寿命。根据量子力学中振动决定于时间的理论，如果一个简并态被激发，则另一简并态被激发的几率不断增加。正如在 4.9.2 中所述，由于核酸分子中最重要的电子激发态是三线态，其寿命约为 10^{-4}～10^{-3}秒（为单线态的一百万倍)，有的至少长达数秒。这对产生电子激发能的共振转移是十分有利的。因而核酸分子能量的共振转移很可能是由一个三线态转移到另一个三线态，以及它们之间的来回转移[11]。

有人认为[12]，蛋白质分子中的各种原子团象摆那样摆动、振动和转动。如果频率相同的两个原子团被例如多肽链偶联起来，那么它们就能进行象偶联振子那样的动作。很可能在同一个蛋白质分子内所有这样的基团都能象偶联振子那样振动，因而给予其中一个基团的能量就可以被其它基团中的任意一个偶联的基团所接受。整个蛋白质分子形成一个统一的系统，而给予它的能量属于整个系统所有，并且可以在这一系统的任何一点表现出来。

11.1.2 激子转移

激子的概念最早由 Frenkel 在 1931 年提出[12]，后来为 Давыдов 所扩展[8]。

固体物理中的激子概念是指：处于激发态的半导体，如果激发的能量不是以使价带中的电子跃迁到导带，则从价带中激发出来的电子只能处在禁带中某些比较稳定的能级上，它和属于同一原子(分子)的空穴相互作用很强，互相束缚在一起形成电子-空穴对，这种激发形式的传播形成一种复式准粒子，称为激子。

Давыдов 把激子看成是晶体中激发的行波束。这种激发不是定域在某个分子上，而是属于分子的聚集体。因此，激子转移是非定域化激发的能量转移形式。

激子转移与半导性能带的电子转移不同。在半导性材料中，正穴与电子沿相反的方向转移，而激子转移则由于正穴与电子共同组成一个激子，因此它们沿相同的方向在晶格中进行传递。激子可以看作是一个电荷对，它们不能分开来移动，直到转移至晶格的适宜点上，电荷偶联丧失，正负电荷就可以分开来在传导带(能带)上进行传导，这就出现半导性导电。所以激子转移与半导性能带的电子转移也密切相关。激子转移通常是在相同分子的电子系统间非常迅速的一种传能方式。由于晶体中的质点安排得非常整齐，电子系统受激后就会完全振动，因而激发能可以由激子转移的方式传递得很远。

以下我们着重介绍Давыдов对于一维均匀分子结构中的激子的研究。

Давыдов 假定，一周期性的一维均匀分子系统沿 z 轴放置，相互间的距离为 R，每一孤立分子的内激发(电子的或振动的)由其能量 ε 和电偶极矩 μ(它与 z 轴的夹角为 θ)标志。并且认为，

分子链足够硬，以致当链转变成激发态时分子间的平衡距离不变。

取链的基态能（当链中无激发分子时）等于零。在一维均匀模型中可以只考虑近邻分子间的相互作用，设链中具有N个分子。N的值足够大，以致可以忽略边缘效应。当第n号分子附近的分子过渡到激发态时，除了考虑确定基态分子间在平衡距离的相互作用外，还须要考虑两个新的相互作用，即：(1) 在分子间呈现出导致从分子n向邻近分子$n\pm1$的过渡激发的共振相互作用，这样的相互作用能可由下式确定

$$J=\mu^2R^{-3}(\cos^2\theta-1) \tag{11.4}$$

(2) 引起原先的相互作用（Van der Waals 及其它相互作用）发生改变的补充相互作用。这种相互作用（$-D$）能量的变化导致引力的增长，因此 $D>0$。当分子n被邻近的分子激发时，由于分子n的相互作用改变，导致在被激发分子的区域内链有变形的趋势。因此，将这种相互作用$-D$称为形变相互作用。

与能量为ε的分子内激发相对应的链的集体激发的能量算符可用下式表示

$$H_{ex}=\sum_n[(\varepsilon-D)B_n^+B_n-J(B_{n+1}^+B_n+B_{n-1}^+B_n)] \tag{11.5}$$

这里的求和是对所有N个分子进行的。算符B_n^+表示分子n被激发，B_n与未激发分子相对应。$(\varepsilon-D)B_n^+B_n$项确定分子n的不包括共振相互作用在内的激发能。算符$B_{n+1}^+B_n$和$B_{n-1}^+B_n$表示分子n向分子$n\pm1$的激发过渡。在(11.5)式中D和J的数值可在未激发链中的分子间的平衡距离R下得出。这个近似的适用范围将在下面确定。

根据量子力学规则，为了确定链集体激发状态的能量，将算符[见(11.5)式]转变成为对角的形式。这可通过么正变换来实现：

$$B_n=\frac{1}{\sqrt{N}}\sum_k A_k\exp(iknR)$$
$$A_k=\frac{1}{\sqrt{N}}\sum_n B_n\exp(-iknR) \tag{11.6}$$

式中k遍及N个分立的、等间隔的数值，$k=\frac{2\pi}{R}v$ $(v=0,\pm1,\pm2\cdots)$，取值范围为

$$-\pi/R<k\leqslant\pi/R \tag{11.7}$$

将(11.6)代入(11.5)，得

$$H_{ex}=\sum_k E(k)A_k^+A_k \tag{11.8}$$

其中

$$E(k)=\varepsilon-D-2J\cos kR \tag{11.9}$$

是链的集体激发态能量，其相应的波函数为

$$\Psi_k(z,t)=\frac{1}{\sqrt{N}}\sum_n e^{i[knR-\omega(k)t]}\varphi(z-nR)$$
$$\hbar\omega(k)=E(k) \tag{11.10}$$

这里，$\varphi(z-nR)$是分子n的激发态波函数，这样的集体激发称为激子。由式(11.10)的函数形式

可见，在形成激子态时，链上全体分子起着同样的作用。换言之，与具有一定k值和能量$E(k)$的激子态相对应的晶体激发是按整个链分布的，而不是集中在一个分子上的。

我们还可将波函数(11.10)改写成如下形式

$$\Psi_k(z,t)=A_k\exp\{i[kz-\omega(k)t]\} \tag{11.11}$$

其中

$$A_k=\frac{1}{\sqrt{N}}\sum_n e^{ik(nR-z)}\varphi(z-nR)$$

函数(11.11)描述$\lambda=2\pi/k$的平面波。然而，我们最感兴趣的是波长大大超过相邻分子间距的激子态。这时$kR\ll 1$，而激子的能量可写成

$$E(k)=E(0)+\frac{\hbar^2k^2}{2m^*} \tag{11.12}$$

式中，第一项$E(0)$是激子的内能

$$E(0)=\varepsilon-D-2J \tag{11.13}$$

第二项相应于有效质量为

$$m^*=\hbar^2/2JR^2 \tag{11.14}$$

的粒子的动能。根据(11.4)，当$\varphi<54°44'$时，有效质量为正，当$\varphi>54°44'$时，有效质量为负。

由平面波(11.11)描述的，具有确定k值和能量$E(k)$的激子态表征定态。在这种状态下，激发能沿整个链分布且不转移。

为了研究激发能沿着链转移的问题，还必须考虑非定态。现在，我们来讨论在时刻t，激发分布在一段链长l_0范围内的状态。这样的状态可用下列波包来描述

$$\Psi(z,t)=\int_{k_0-\Delta k}^{k_0+\Delta k}A(k)\exp\{i[kz-\omega(k)\}\mathrm{d}k \tag{11.15}$$

式中$\Delta k=\pi/2l_0$。从(11.15)得，激发以群速度

$$v=\left[\frac{\partial\omega(k)}{\partial k}\right]_{k=k_0}=\frac{\hbar k_0}{m^*} \tag{11.16}$$

沿链传播，其中$\hbar k_0$是处于态(11.15)的激发的平均动量。

于是，在时刻t于l_0一段链上的激发态按(11.16)式所确定的速度沿着分子链传播，这时能量$E(k_0)$沿着链转移。共振相互作用$|J|$愈大，有效质量$|m^*|$就愈小。相应地，当k_0值给定时，激子运动的速度也就愈大。

当激子使能量沿着分子链转移时，应当考虑两个重要的情况：

(1) 如果在时刻t被激发所围的链长由l_0确定，则在时刻$t+\tau$时它将是

$$l_\tau=\sqrt{l_0^2+(\hbar\tau/m^*l_0)^2}$$

$$l_0=\frac{2\Delta k}{\pi} \tag{11.17}$$

所谓波包随着时间的增加而“蔓延”，当$\tau>m^*l_0/k$时，波包的宽度按速度$\hbar/m^*l_0$正比于时间而增长。激子的有效质量越小，扩散越明显。

波包随时间的蔓延是由波包(11.15) 所描述的激子状态不太稳定的标志之一。在这样的状态中不管是激子的能量和动量都不具有严格确定的值,而具有确定k和$E(k)$值的状态，则按系统均匀分布且不转移能量。

(2) 当激子自身运动时它被声子散射，也就是在链中靠近它们的平衡位置附近的分子振动上散射,当发射(吸收)声子而散射时,它们将损失(获得)能量。

激子在声子上的散射过程取决于形变相互作用能$-D$对分子离开它的平衡位置的位移u_n的依赖关系。分子内激发与分子位移的这种联系可以写成

$$H_{\text{int}}=x\sum_n(u_{n+1}-u_{n-1})B_n^+B_n \tag{11.18}$$

式中$X=\left(\frac{\partial D}{\partial u_n}\right)_{u_n=0}$是激子和声子的耦合参数。

在低温时的激子散射中只有产生声子的激子能量损失过程起主要作用。这样的过程只有当激子速度超过纵向声波的速度时才有可能。如果w是分子链的纵向弹性参数，M是单个分子的质量，则纵向声波的速度由下式确定

$$v_a=R\sqrt{w/M} \tag{11.19}$$

因此，发射声子时制动激子的条件由下面的不等式确定

$$\frac{\hbar k}{m^*}>v_a\equiv R\sqrt{w/M} \tag{11.20}$$

可见,在一维分子链中这样的理论仅对于那些具有满足不等式(11.20)的波数k的激子才是正确的。事实上,当激子的速度超过纵声波的速度时,在激发通过的时间内分子才来不及从它自己的平衡位置R处离开。

如果不等式(11.20)不满足,则对于低温声子发射来说，激子的寿命τ_{ex}将反比于进入算符(11.18)的参数X的平方以及分子在平衡位置附近作零点振动的振幅平方,后者由下式确定

$$\langle u^2\rangle_0=\hbar/2\sqrt{wM} \tag{11.21}$$

这样,

$$\tau_{ph}=[\langle u^2\rangle_0X^2]^{-1}=\frac{2\sqrt{wM}}{\hbar X^2} \tag{11.22}$$

如果温度不等于零,则式(11.21)改写为

$$\langle u^2\rangle_T=\frac{\hbar}{2\sqrt{Mw}}\text{cth}\frac{\hbar\sqrt{w/M}}{2\Theta}$$

式中Θ是绝对温度,以能量单位表示。

在二维和三维周期性分子系统中，分子激子型的集体激发状态也是可能产生的。这些系统中的激子不是用波数K来标志,而是用波矢K和能量$E(K)$标志的。

分子系统中的激子是由电磁波——光子产生的。如果光子的能量为$\hbar\omega$，其波数$|\mathbf{Q}|=\omega/c$(这里c是光速)，则在同时遵循能量守恒定律

$$\hbar\omega=E(K) \tag{11.23}$$

和动量守恒定律

$$\hbar Q=\hbar K \tag{11.23'}$$

的条件下，即使声子不参与，光子转变为激子（或相反）也是能实现的。

在光谱范围内，辐射的波长远远大于分子间距R，$QR\ll1$。这时，只有$k\approx0$ 的激子与辐射产生相互作用。因此等式(11.23)可以用近似等式 $\hbar\omega=E(0)$ 代替。这样与激子产生相互作用的辐射的波长可由下式确定

$$\lambda=\frac{2\pi c}{\omega}=2\pi\hbar c/E(0)$$

对于光子辐射激子的寿命 τ_{ex} 已由Агранович[13] 计算过，由下式给出

$$\tau_{ex}=\frac{2E(0)R}{3\pi c\hbar}\tau_0=\frac{4R}{3\lambda}\tau_0 \tag{11.24}$$

式中 $\tau_0=3\hbar^4c^3/4E^3(0)\mu^2$ 是孤立分子具有能量$E(0)$的激发寿命。

由于$R/\lambda\approx10^{-3}$，所以在一维分子链中对于光子辐射的激子寿命比自由分子的激发寿命短一千倍。量 $1/\tau_{ex}$ 确定由光子激发激子的几率。按照(11.24) 式，在一维分子链中光子激发的几率超过自由分子的激发几率 $1/\tau_0$ 一千倍。

与激发分子的最低电子态相应的激子由可见光和紫外光激发，与分子内振动，例如蛋白质分子中的酰胺 I 振动（波数为 1660 cm^{-1}）相应的激子可由红外辐射激发。Moffitt[14]曾经就 α-螺旋蛋白质中酰胺基的 $\pi\pi^*$190 nm 跃迁的激子分裂作过理论计算，结果与 Давыдов 的理论相符，这个吸收带分裂为两部分，一个沿螺旋轴偏振，另一个垂直于螺旋轴。第一个带的频率是 2700 cm^{-1}，低于第二个带，与实验极好地一致。

Reid 曾指出，在生物体内普遍存在大分子的有序但非均质的系统，激子转移可能有相当的重要性。DNA 分子本身具有高度有序的结构，它周围还存在着类似石英型结晶的液晶态有序水结构，而且 DNA 的介电常数高达 120 000 以上，这就为激子转移创造了条件，并且可以使激子保存较久。

在生物现象中，激子还在光合作用中起着基本的作用。当吸收光时，叶绿素分子可形成激子。当激子与声子相互作用时很快地损失能量，而将其转移至光合作用的中心。

11.1.3 孤子转移[3]

Давыдов 等认为，生物系统中，ATP 水解释放出的能量（0.43 eV 或 10 kcal/mol）不足以激发分子中的电子态，而对蛋白质分子中的分离原子集团的振动激发却具有重要意义。这种振动激发（产生位移）与极化激发（产生电偶极矩）不可分离地联结在一起。这种激发同样可以导致整个分子链的集体激发。他们从上述激发下分子链的 Hamilton 形式出发，运用半经典的量子力学方法，建立了一个非线性的 Давыдов 方程组，获得了它的孤子群，并论证这种孤子的稳定性，从而说明了孤子这种集体携带能量沿一维分子链传输的高效性。

我们知道，对于某些类型的非线性波动方程，可以具有稳定的行波解，其形状是铃状脉冲的孤立波，它的特点是在运动中能够保持波形不变，而且与其它孤立波发生碰撞后，仍然不改变波

形和速度，具有这些特点（如定域特点）的孤立波就称为孤子。

11.1.3.1 一维弹性分子结构中的孤子

Давыдов 等认为，如果激发沿分子系统的移动速度小于声速，则该运动将在分子的激发范围内伴有系统的局域形变。因此，当计算集体激发能时应当考虑分子从平衡位置的位移能[15,16]。这样的复杂激发状态的能量算符应该包括内激发能量算符(11.5)，分子离开平衡位置的位移 u_n 的能量算符 H_{ph} 和内激发与位移的相互作用算符(11.18)。

在简谐近似下分子位移的能量算符可写成

$$H_{ph}=\frac{1}{2}\sum_n\left[\frac{1}{M}P_n^2+w(u_n-u_{n-1})^2\right] \tag{11.25}$$

式中，M是分子的质量，w 是线性链的弹性系数，P_n 是与位移算符 u_n 共轭的正则动量算符。

这样，集体激发的能量 Hamilton就包含三个相加项

$$H_{sol}=H_{ex}+H_{ph}+H_{int} \tag{11.26}$$

我们可将激发态写成如下函数形式

$$|\Psi(t)\rangle=\sum_n \alpha n(t)e^{\sigma(t)}B_n^+|0\rangle \tag{11.27}$$

式中函数$|0\rangle$描述链中全体分子处于平衡位置的基态

$$\sigma(t)\equiv-\frac{i}{\hbar}\sum_n\left[\beta_n(t)P_n-\pi_n(t)u_n\right] \tag{11.28}$$

由函数(11.27)的正交归一条件得

$$\sum_n|\alpha_n(t)|^2=1 \tag{11.29}$$

由泛函

$$\Phi(\{\alpha_n,\beta_n,\pi_n\})\equiv\langle\Psi(t)|H|\Psi(t)\rangle \tag{11.30}$$

取极小值的变分法可得复函数 $\alpha_n(t)$和实函数 $\beta_n(t)$、$\pi_n(t)$，这时量$|\alpha_n|^2$ 确定沿分子链的分子内激发分布的几率。

利用(11.26)—(11.28)式可得

$$\langle\Psi(t)|u_n|\Psi(t)\rangle=\beta_n(t)$$
$$\langle\Psi(t)|P_n|\Psi(t)\rangle=\pi_n(t)$$

由此可见，函数 $\beta_n(t)$和$\pi_n(t)$表征了由函数(11.27)确定的状态中，分子位移和动量的平均值。

泛函(11.30)取极小值的条件归结为求解联合方程组

$$\left.\begin{aligned}i\hbar\frac{\partial\alpha_n}{\partial t}&=\frac{\partial\Phi}{\partial\alpha_n^*}\\&=[\varepsilon_0+w+\chi(\beta_{n+1}-\beta_{n-1})]\alpha_n-J(\alpha_{n+1}+\alpha_{n-1})\\\frac{\partial^2\beta_n}{\partial t^2}+\frac{w}{M}(2\beta_n-\beta_{n+1}-\beta_{n-1})&=\frac{\chi}{M}(|\alpha_{n+1}|^2-|\alpha_{n-1}|^2)\end{aligned}\right\} \tag{11.31}$$

这时

$$\varepsilon_0 = \varepsilon - D\,,\ \pi_n = M\frac{\partial \beta_n}{\partial t}$$

$$W = \frac{1}{2}\sum_n \left[M\left(\frac{\partial \beta_n}{\partial t}\right)^2 + w(\beta_n - \beta_{n-1})^2 \right]$$

然而，我们感兴趣的仅仅是这样的激发，在这种激发下，函数$\alpha_n(t)$、$\beta_n(t)$对于分子的号码的依赖是平滑的。这样，可以用没有量纲的自变量$\xi = Z/R$的连续函数$\alpha(\xi,t)$，$\beta(\xi,t)$来代替这些函数，以使下列等式成立

$$\alpha(n,t) = \alpha_n(t), \quad \beta(n,t) = \beta_n(t)$$

将近似方程(11.31)变成为下列微分方程：

$$\left.\begin{aligned} &\left[i\hbar\frac{\partial}{\partial t} - \Lambda + J\frac{\partial^2}{\partial \xi^2} - 2\chi\frac{\partial \beta(\xi,t)}{\partial \xi}\right]\alpha(\xi,t) = 0 \\ &\left[\frac{\partial^2}{\partial t^2} - v_\alpha^2\frac{\partial^2}{\partial \xi^2}\right]\beta(\xi,t) - \frac{2\chi}{M}\frac{\partial}{\partial \xi}|\alpha(\xi,t)^2| = 0 \end{aligned}\right\} \tag{11.32}$$

式中$\Lambda = \varepsilon_0 + W - 2J$；$v_\alpha^2 = w/M$。而$V_\alpha = v_\alpha R$的值决定纵声波在分子链中的速度。替代表征个别分子在未激发链中对于平衡位置位移的函数$\beta(\xi,t)$可引入表征邻近分子间距R相对减小的函数

$$\rho(\xi,t) = -\frac{\partial \beta(\xi,t)}{\partial \xi} \tag{11.33}$$

这样，方程(11.32)成为

$$\left[i\hbar\frac{\partial}{\partial t} - \Lambda + J\frac{\partial^2}{\partial \xi^2} + 2\chi\rho(\xi,t)\right]\alpha(\xi,t) = 0 \tag{11.34}$$

$$\left[\frac{\partial^2}{\partial t^2} - v_\alpha^2\frac{\partial^2}{\partial \xi^2}\right]\rho(\xi,t) + \frac{2\chi}{M}\frac{\partial^2}{\partial \xi^2}|\alpha(\xi,t)|^2 = 0 \tag{11.35}$$

我们将寻求方程组(11.34)，(11.35)的，以速度$V = vR$沿着链传播的激发态的解，也就是设

$$\rho(\xi,t) = \rho(\xi - vt)$$

$$\alpha(\xi,t) = \Phi(\xi - vt)e^{iv(\xi,t)}$$

这样，从(11.35)得

$$\rho(\xi,t) = \frac{2\chi|\alpha(\xi,t)|}{w(1-S^2)} \tag{11.36}$$

式中S是激发传播速度与纵声波速度之比

$$S = \frac{v}{v_\alpha} = \frac{V}{V_\alpha} < 1$$

将(11.36)代入(11.34)，我们就得到确定函数$\alpha(\xi,t)$的非线性 Schrödinger 方程

$$\left\{i\hbar\frac{\partial}{\partial t} - \Lambda + J\frac{\partial^2}{\partial \xi^2} + G|\alpha(\xi,t)|^2\right\}\alpha(\xi,t) = 0 \tag{11.37}$$

式中非线性参数

$$G = \frac{4\chi^2}{w(1-S^2)} \tag{11.38}$$

归一化条件(11.29)在近似下具有如下形式

$$\int_{-\infty}^{\infty}|\alpha(\xi,t)|^2=1$$

方程(11.37)的归一化的精确解由函数

$$\alpha(\xi,t)=\frac{\sqrt{\mu}\exp\left\{i\left[\frac{\hbar v}{2J}(\xi-\xi_0)-E_v t/\hbar\right]\right\}}{\sqrt{2}\,c\hbar\,[\mu(\xi-\xi_0-vt)]} \tag{11.39}$$

给出。将该值代入(11.36),就可找到分子间距相对减小的函数

$$\rho(\xi,t)=\frac{\chi_\mu}{w(1-S^2)c\hbar^2[\mu(\xi-\xi_0-vt)]} \tag{11.40}$$

模的平方

$$|\alpha(\xi,t)|^2=\frac{\mu}{2c\hbar^2[\mu(\xi-\xi_0-vt)]} \tag{11.41}$$

确定单个分子沿着链激发的几率分布。ξ_0 表示 $t=0$ 时的最大形变和激发的最大几率位置。根据(11.40)和(11.41),形变和激发以速度 $V=vR$ 沿着链传播。表达式(11.39)—(11.41)中的参数

$$\mu=\frac{\chi^2}{wJ(1-S^2)}$$

确定表征链的激发部分

$$\Delta Z=\pi R/\mu_0 \tag{11.42}$$

由函数(11.39)—(11.42)描述的激发称为孤子,以速度 $V=vR$ 运动的孤子传递能量

$$E_V=E_0+\frac{1}{2}m_{sol}v^2$$

式中 E_0 是孤子的内能

$$E_0=\varepsilon-D-2J-\frac{\chi^4}{3w^2J},$$

m_{sol} 是孤子的有效质量

$$m_{sol}=\frac{\hbar^2}{2JR^2}+\frac{4\chi^4\left(1+\frac{3}{2}S^2-\frac{1}{2}S^4\right)}{3w^2JV_a^2(1-S^2)^3}$$

当所有 E_V 的值有限时,孤子的速度总是小于分子链中纵声波的速度,即 $S\ll1$。在软链(小的 w 值)中以及当内激发与分子位移紧密联系时(χ 的值大)激子的有效质量很大。因此,甚至当运动速度较小时孤子的动能仍可能是很大的。

当存在孤子激发时分子链局域地形变,发射能量后该局域形变应当消失。局域形变能耗散成为链中分子振动的热能。由于这个能量并不辐射,所以辐射的极大值与频率

$$w=\frac{1}{\hbar}E_V-A \tag{11.43}$$

相对应,这里 A 是 Stokes 位移

$$A=\frac{2\pi\chi^4}{3\hbar w^2 J} \tag{11.44}$$

在软链中以及当分子内激发与分子位移有紧密联系(这种联系由参数 χ 确定)时,A 的值就大。

根据 Давыдов[15]的计算,孤子的发射寿命由下式确定:

$$\frac{1}{\tau_{sol}}=\frac{2\pi^2 R}{3\lambda\mu c h^2\left(\frac{2\pi\hbar v}{2\mu J}\right)}\frac{1}{\tau_{ex}} \tag{11.45}$$

式中$\frac{1}{\tau_{ex}}$和$\frac{1}{\tau_{sol}}$是激子和孤子的辐射几率。静止孤子的辐射几率(11.45)具有极大值。

根据(11.42)$\Delta Z=\pi R/\mu$,如果考虑到分子链中被激发所包围的区域,则当 $v=0$ 时可以写成

$$\frac{1}{\tau_{sol}}=\frac{2\pi\Delta Z}{3\lambda}\cdot\frac{1}{\tau_{ex}},\quad \Delta Z=\frac{\pi RwJ}{\chi^2} \tag{11.46}$$

因此,对于软链来说,当Stokes 位移比较大,$\Delta Z\ll\lambda$,就有不等式

$$\tau_{sol}^{-1}\ll\tau_{ex}^{-1} \tag{11.47}$$

只有在局域作用下才能激发孤子。化学反应与这种局域作用相对应。化学反应发生于分子链的一端时,所释放出的能量导致产生孤子,它将沿分子链移动。

11.1.3.2 激子与孤子性质的比较

正如前面所述,在一维分子系统中可以产生两种类型的集体激发——激子与孤子。

激子在链中移动的速度大于纵声波在链中的速度,因此它来不及引起链的局域形变。孤子也是一种集体激发,它在链中移动的速度小于纵波的速度,这种激发是分子内激发与形变激发的组合。孤子的内能小于组成它的各种激发的能量之和,这是孤子具有大的稳定性的基本原因之一。由于孤子的运动伴有链的形变,所以它的有效质量远大于激子的有效质量。在软链中以及当内激发与分子位移有强耦合时,这个差别就特别大。激子仅携带分子内激发能量,它的有效质量反比于谐振相互作用,当运动时它被制止,并发射声子。然而,由于孤子有大的有效质量,其运动速度小于声速,因此它不将自己的能量耗费于发射声子。

由于孤子极大的稳定性使其在运动时因不发射声子因而能量也无损耗,这些都类似于金属的超导电性。

由 501 页表 11-1 可见,孤子的内能比激子的内能小 $\chi^4/3w^2J$,因此形成孤子在能量上比形成激子较有利。

当晶格硬度增加($w\to\infty$),共振相互作用增加和内激发与分子位移耦合减小时,孤子和激子在性质上的差别就会越来越小。孤子的有效质量就接近于激子的有效质量。孤子态中被激发所包围的范围就增加。

11.1.3.3 α-螺旋蛋白质中分子结合能沿一维分子链传输的孤子运动[17]

1973 年以来,Давыдов 等对具有由弱相互作用的周期性重复的原子集团组成的准一维 α-螺旋蛋白质分子系统进行了深入研究(参见图 11-1)。他们在假定 α-螺旋蛋白质分子三条链中每一条链都独立激发的基础上,研究了与多肽中酰胺-I 振动相应的分子集体激发。这时问题可

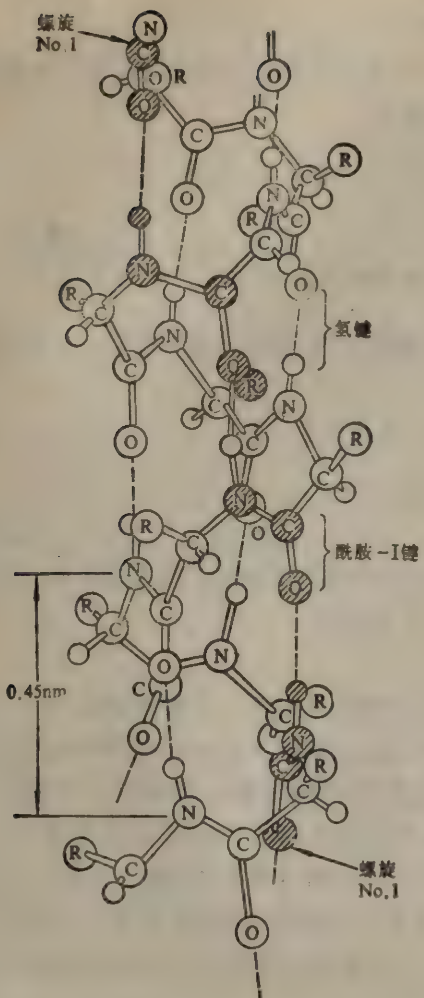


图 11-1　α 螺线蛋白质的结构

归结为弱耦合肽基的一维分子系统，这样的激发前面已经讨论过。Давыдов 等的研究进一步表明，如果由弱相互作用的周期性重复的原子集团组成的准一维 α-螺旋蛋白质分子系统受到光的照射或经化学反应等作用后，在 α-螺旋蛋白质中，通过相邻肽基之间的偶极-偶极相互作用，或在 ATP 分子的水解作用中所释放出来的能量，便可在其中引起一种集体激发。由于 α-螺旋蛋白质分子结构的特点（如图 11-1），这种集体激发并不局限于一处，而是可携带能量沿一维 α-螺旋蛋白质分子轴线传输。由此他们推算出了结合能的传输和肽基间距变化所满足的非线性的Давыдов 方程组（见文献[3]）。但此方程组还未能很好地反映 α-螺旋蛋白质分子所具有的所有特点，因此，后来人们在修改中，全面地考虑了这些对称性和纵向氢键的线性恢复力，并引起了十个反映各种偶极-偶极相互作用的附加参数[18,19]。则修正后一个完整的 Давыдов 方程组具有如下形式

$$i\frac{\mathrm{d}A_{m,i}}{\mathrm{d}\tau}+J(A_{m+1,i}-2A_{m,i}+A_{m-1,i})-\kappa_1(B_{m+1,i}-B_{m-1,i})A_{m,i}$$
$$=LF_L-2JA_{m,i}+\{\kappa_2[B_{m-1,i}A_{m-1,i}-B_{m+1,i}A_{m+1,i}+B_{m,i}(A_{m+1,i}+A_{m-1,i})]+NF_N+PF_P+QF_Q+RF_R+SF_S+TF_T+UF_U+VF_V+XF_X+ZF_Z\} \tag{11.48}$$

$$\frac{\mathrm{d}^2B_{m,i}}{\mathrm{d}\tau^2}-(B_{m+1,i}-2B_{m,i}+B_{m-1,i})$$
$$=K_3(|A_{m+1,i}|^2-|A_{m-1,i}|^2)+\{K_4[A^*_{m,i}(A_{m+1,i}-A_{m-1,i})+A_{m,i}(A^*_{m+1,i}-A^*_{m-1,i})]\} \tag{11.49}$$

这里，$A_{m,i}$ 是结合能的几率密度，$B_{m,i}$ 表示一个酰胺-I 键的纵向传移。常数（J,L,N,P,Q,R,S,T,U,V,X,Z）分别表示酰胺-I 键对之间的各种偶极-偶极相互作用能系数。在计算单位下，它们的数值分别为{0.145，0.231，0.073，0.034，0.019，0.012，0.0089，0.0073，0.0037，0.0030，0.0022，0.0017}。m 表示沿螺线的元胞编号，i 表示三个肽基链的标号，即 $i=1,2,3$。符号 $\{F_L,F_N,F_P,F_Q,F_R,\cdots,F_Z\}$ 是一种缩写，它们有确切的定义。例如，F_L 表示横向相邻的酰

表 11-1

参　　数	激　　子	孤　　子
内　　能	$E^0_{ex}=\varepsilon-D-2J$	$E^0_{sol}=E^0_{ex}-x^4/3w^2J$
有　效　质　量	$m_{ex}=\hbar^2/2JR^2$	$m_{sol}=m_{ex}+\frac{4x^4M}{3w^2JR^2}$
激　发　所　包 围　的　线　度	$l_t=\sqrt{l_c^2+\left(\frac{\hbar t}{m_{ex}l_0}\right)^2}$ $l_0=2\Delta k\ \pi$	$l=\pi RwJ/x^2$
产生波长为λ的辐射时的发射寿命	$\tau_{ex}=\frac{4R}{3\lambda}\tau_0$	$\tau_{sol}=\frac{3\lambda x^2}{2\pi^2RwJ}\tau_c$
按频率的 Stokes 位移	~ 0	$\frac{2\pi x^4}{3w\hbar J}$

胺-I 键之间的偶极-偶极相互作用。它被定义为：

$$
\begin{aligned}
iA^0{}_{m,1}&=L(A_{m,2}+A_{m-1,3})+\text{etc}\\
iA^0{}_{m,2}&=L(A_{m,3}+A_{m,1})+\text{etc}\\
iA^0{}_{m,3}&=L(A_{m,2}+A_{m+1,1})+\text{etc}
\end{aligned}
\tag{11.50}
$$

其它符号如$F_N\cdots F_Z$也有类似的定义。

常数$\{k_1,k_2,k_3,k_4\}$分别表示结合能色散和纵向声波之间的各种非线性耦合系数，它们被定义为：

$$
\left.
\begin{aligned}
k_1&=\frac{10^{-11}}{\hbar}\sqrt{\frac{M}{w}}\chi_1, &\qquad k_2&=\frac{10^{-11}}{\hbar}\sqrt{\frac{M}{w}}\chi_2,\\
k_3&=\frac{10^{11}}{w}\chi_1, & k_4&=\frac{10^{11}}{w}\chi_2
\end{aligned}
\right\}
\tag{11.51}
$$

其中M表示一个单位元胞的质量的$\frac{1}{3}$。w是氢键的线性力常数。χ_1表示在一个相邻氢键每单位伸长下，在酰胺-I键中的结合能的变化。χ_2表示在单位微分位移存在下，所引起的最近邻纵向偶极-偶极相互作用项J的变化。对于时τ以$\sqrt{\frac{M}{w}}=0.99\times10^{-12}$s作测量单位，$B_{m,i}$以$10^{-11}m$作为测量单位。

Scott 等[19]用数值计算的方法研究了这组非线性方程的解。计算结果表明：Давыдов 孤子确实能在α-螺旋蛋白质中形成。并由此求出了这种孤子能出现的判定参数χ_1的阈值为 0.34×10^{-10}Newton。

我国科学工作者庞小峰近年来用解析方法研究了这组修正的 Давыдов方程组的孤子解(详见文献[17])。

从Давыдов、Scott和庞小峰等的研究可以看出，在α-螺旋蛋白质中，其分子结合能 确实是

以孤子形式沿其螺旋轴线进行传播的。这种孤子所具有的确定能量和质量与酰胺-I键之间的各种偶极-偶极相互作用以及孤子传播速度有关。孤子携带这种能量在α-螺旋蛋白质中传输时，必然会引起一些物理现象和生理过程的出现。

Давыдов等的研究，揭示了孤子在传递生物化学能通过生物系统方面所能起到的某些重要作用。例如，在分子水平上，孤子可能沿着一个α-螺旋蛋白质分子传递肽基的C=O特征振动的能量而不激发热运动，因而极少耗损。Давыдов还应用这同样的概念在分子水平上解释了动物肌肉的收缩机理[3]，而 Scott 等则将其应用于研究细胞膜中的非线性电动力学[20]和神经系统等中的孤子运动问题。

总之，从生物学的观点看，例如当某一个细胞受到某种"激发"，则这种"激发"也必然会通过细胞间的相互关联而传递开来，从而可能遍及整个细胞链甚至更大范围。同样，某一大分子受到某种"激发"，这一"激发"也会遍及整个分子链；分子内的电子系统也是如此。对于那些在结构上具有一定规则排列(或具有准周期性)的细胞系统，大分子链或大分子本身来说，这种集体性的"激发"更为显著。

生物体的某些生理现象：植物叶绿素分子接受光的照射而获得的能量，可以很快地转移至光合作用中心而贮存起来。再如我们的感官(眼、耳、舌、鼻和皮肤等)，其所接受的外界刺激可以各不相同(机械、辐射、化学，···)，受体本身的生理结构和接受刺激后的生物物理、生物化学过程各不相同，但最终总要把这些因受到外来刺激而引起的"激发"转变为电信号，而后通过各自不同的神经纤维将这种"激发"传送至中枢，而中枢神经组织作出的反应也通过神经传至各个组织和器官，这种激发的传播是几乎不衰减、不畸变的。从下一节开始，我们将讨论通过其它形式(如借助于电子、质子等)进行的能量转移。

我们知道，固体物理中引入的准粒子有好几种，它们是由不同的相互作用类型，不同的集体相关运动方式而建立的，其中如激子和孤子等几种准粒子完全可以用于生物系统。

11.2 生物系统中的电子转移

前面叙述了激发能的共振转移、激子转移和孤子转移，下面将继续讨论其它形式的能量转移。

许多生物现象——在线粒体、细胞色素和神经冲动的通过以及其它系统中电子传递链的作用——跟电荷(电子和离子)的转移有关，与中性激发(激子、孤子、声子)的能量转移不同，电荷的转移只有当存在电场或浓度梯度时才有可能。

已知电场强度抵消电势梯度φ。由于静电相互作用势的特点是：当电子转移时释放或吸收的能量取决于电势差而与电子转移的路径无关。因而具有负电荷 $-e$的电子从电势为 φ_1的一点到达电势为φ_2的一点的位移是与释放能量$e(\varphi_2-\varphi_1)$相联系的(如果 $\varphi_2>\varphi_1$)。反之，如果向相反方向运动则要消耗相应的能量。

就电子转移而言，在生物体内大致分为两种：一种是关于酶反应时酶与底物之间，或酶-辅酶之间的电子转移等为代表的电子非定域化。所谓电荷转移，全都是这样的电子转移。因为电

子是通过轨道的重叠而转移，所以在比较近的距离内发生。另一种是在生物体内的氧化-还原反应中，电子通过一些中间电子传递体顺次传递（参见图10-38），最后交给受体的这样一种电子转移，其机制至今还不十分清楚[21]。

迄今为止，有关激发能借助于电子转移而进行能量转移的研究已经很广泛，几种重要的生物大分子如蛋白质与核酸，以及几种重要的光受体中的物质如叶绿素与视觉色素（包括黑色素与视黄醛）等都进行了一些研究。然而，尽管在电子转移方面已经积累了不少资料，但目前还不能总结出通过电子转移的能量转移规律。

11.2.1 络合物电荷转移[22,23]

在生物学过程中意义更为普遍的传能方式可能是络合物的电荷转移，它在生物能力学中的意义正是Szent-Györgyi所著《亚分子生物学导论》一书的基本思想[22]。电荷转移表明组成络合物的一个分子D(电子供体)的电子可以享用另一个分子A(电子受体)的轨道，因为电荷迁移络合物的形成，与组成络合物的分子的供、受电子的能力有关。

在大分子内部的基团之间，也可以产生电子的供受，从而表现类似络合物电荷转移的传能形式。电子从一个基团的最高满带自发地转移到另一个基团的最低导带，这种机制在生物体中可能较为普遍。如果大分子的结构和能量关系符合于形成电荷迁移集合体的要求，那么激发能就可以通过电子在这些分子所组成的骨架中进行转移。例如，在DNA分子的双螺旋结构中，以特殊定向排列原子团是一个很重要的事实，其中的嘌呤和嘧啶基团可能发生电荷转移。在嘌呤和嘧啶的平行排列架中，实现电荷转移的条件比溶液中优越，而且电荷转移过程可沿这种分子架反复进行。

电荷转移络合物是 Mulliken 最早提出的，他运用价键理论来描述这些络合物的结合和能级[24]。Mulliken 理论是一种价键理论，该理论指出，电荷转移络合物(D·A) 的状态由 D与 A 的相互作用的非键态(D,A) 和配键态($D^+—A^-$)的组态线性组合而成。

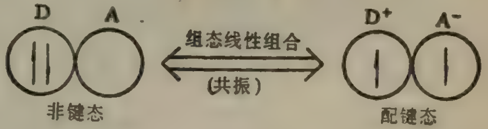


图 11-2 电荷转移络合物中的非键态(D, A)和配键态($D^+—A^-$)

若用量子力学的符号表示，上述组态线性组合过程可写为式(11.52)，也可用图 11-2 表示。

$$\psi(D\cdot A)=a\psi_0(D,A)+b\psi_1(D^+—A^-) \tag{11.52}$$

其中a和b是组合系数。

如何决定$\psi(D\cdot A)$，即求出a和b。将(11.52)式代入Schrödinger 方程，并根据变分原理得体系能量E的两个根E_N和E_E，

$$\left.\begin{aligned} E_N&=\left\{\frac{1}{2}(E_0+E_1)-S_{01}H_{01}-\left[\left(\frac{\Delta}{2}\right)^2+\beta_0\beta_1\right]^{1/2}\right\}\Big/(1-S^2_{01})\\ E_E&=\left\{\frac{1}{2}(E_0+E_1)-S_{01}H_{01}+\left[\left(\frac{\Delta}{2}\right)^2+\beta_0\beta_1\right]^{1/2}\right\}\Big/(1-S^2_{01})\end{aligned}\right\} \tag{11.53}$$

E_N和E_E各为电子供受络合物的基态和激发态的能量。式(11.53)中

$$\Delta \equiv E_1 - E_0$$

（通常$\Delta>0$，但有时也可能$\Delta<0$） (11.54)

$$\beta_0 \equiv H_{01} - S_{01}E_{01}$$

$$\beta_1 \equiv H_{01} - S_{01}E_1 (\beta_0<0, \beta_1<0) \tag{11.55}$$

其中

$$E_0 \equiv H_{00} = \int \psi_0{}^+ \hat{H} \psi_0 \mathrm{d}\tau;$$

$$E_1 \equiv H_{11} = \int \psi_1{}^+ \hat{H} \psi_1 \mathrm{d}\tau$$

$$H_{01} = H_{10} = \int \psi_0{}^+ \hat{H} \psi_1 \mathrm{d}\tau$$

$$= \int \psi_1{}^+ \hat{H} \psi_0 \mathrm{d}\tau$$

S则为重叠积分。有了式(11.53)，则不难求出两套线性组合系数(a_N, b_N)及(a_E, b_E)，从而求得基态和激发态的态函数

$$\begin{cases} \psi_N = (\psi_0 + \rho_N \psi_1)/(1 + 2\rho_N S_{01} + \rho_N{}^2)^{1/2} \\ \psi_E = (\psi_1 - \rho_E \psi_0)/(1 - 2\rho_E S_{01} + \rho_E{}^2)^{1/2} \end{cases} \tag{11.56}$$

其中

$$\begin{cases} \rho_N = b_N/a_N = -(H_{01} - S_{01}E_N)/(E_1 - E_N) \\ \rho_E = -a_E/b_E = (H_{01} - S_{01}E_E)/(E_0 - E_E) \end{cases} \tag{11.57}$$

根据此理论，电荷转移络合物的电荷转移谱带的频率为

$$h\nu_{CT} = E_E - E_N = 2\left[\left(\frac{\Delta}{2}\right)^2 + \beta_0\beta_1\right]^{1/2} \Big/ (1 - S_{01}^2) \tag{11.58}$$

在 Mulliken 提出电荷转移络合物的价键理论后，便有人企图用分子轨道理论来描述供体与受体的相互作用。根据所采用的分子轨道的基函数及推导中的近似程度，文献中已有几种关于电荷转移络合物的分子轨道理论[25]。其中，Flurry 提出的分子轨道理论有可能广泛地用来计算实际的体系(见文献[26])。

除了 Mulliken 和 Flurry 等的理论外，还有用整体分子轨道处理的量子化学方法来进行计算的。假如不涉及 D 和 A 在 D·A 中所起的作用如何，则从纯粹的计算结果来看，用整体分子轨道的量子化学方法进行计算将获得更准确的结果。在计算中可以用 HMO、EHMO、CNDO、INDO 或 ab initio 等任一方法。当运用这些方法于电荷转移络合物的分子时，其手续与运用于其它经典的分子的情况完全一样。

络合物电荷转移在激发能的传递上有许多特点。归纳起来大致有：(1) 电荷转移可分为自发的和光致的两种。前者又叫强电荷转移，后者又叫弱电荷转移。强电荷转移不需要光能激发就可以自发地产生，因而并不需要很大的激活能就可形成激发能级。络合物电荷转移传能的大小与生物体所能利用的能量子(例如 ATP 高能磷酸键的 10～12 kcal)大小十分一致。(2) 络合物电荷转移使电子从一物质向另一物质转移而能量并无多大损失，因为它并不引起分子结构

或化学键的重新排列。(3) 电荷转移络合物是介于有规则闭壳分子和自由基之间的一种物质状态。络合物经电荷转移后,如果离解成电子供体和电子受体的单分子,就可以形成两个具有非偶电子的自由基[22]。

正是由于络合物电荷转移具有这些特点,因而被认为是活体内转移电子 激发 能最有研究前途的机制之一。Szent-Györgyi 在«亚分子生物学导论»一书中,就特别强调这 种传 能 机制对生物学的重要性。他甚至说,“从分子的基态直接跃迁到同一分子的激发态,对于不 是光致的过程来说,似乎不太可能,因此我曾对激发态和电子激发能失去信心。但电荷传递又帮助我恢复了这种信心。”[22]

11.2.2 电子的隧道转移[21]

已经知道,在细胞呼吸过程中有一系列的电子传递体在起作用(参见 图 10-38),图 11-3 表示在一个电子传递体中的势垒和势阱。

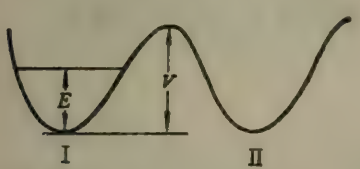


图 11-3

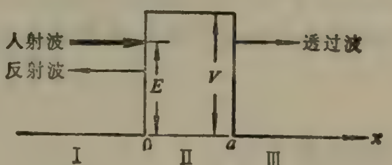


图 11-4 对势垒入射粒子(波动)的穿过和反射

对于线粒体而言,其电子传递体为细胞色素的铁卟啉(见第 10 章),核 黄 素 蛋 白 的 辅基,NADH,NADPH,非血红蛋白的铁离子,细胞色素氧化酶的铜离子;对于光合作用系统则是细胞色素,质体兰素的活性中心,铁氧还原蛋白等。在电子的隧道转移中如下反应起着重要的作用

$$Fe^{3+} + e \longrightarrow Fe^{2+},$$

$$Fe^{2+} - e \longrightarrow Fe^{3+}$$

按照量子力学的观点,电子(在图 11-4 中)从 I 转移至 II 不是越过势垒 V 而是通过“隧道效应”进行的。为使问题简化,我们将图 11-3 的势换成图 11-4 的形状。

在 $x<0, x>a$ 时,$V(x)=0$; 在 $0<x<a$ 时,$V(x)=V$

当具有能量 E 的粒子和能量 $V(>E)$的势垒碰撞时,大部分弹回,而一部分穿过势垒达到 III 的区域。I,II,III 各区域粒子的波动函数可写成:

$$\text{I} \quad x \leqq 0 \quad \varphi(x) = Ae^{i\kappa x} + Be^{-i\kappa x} \tag{11.59}$$

$$\text{II} \quad a > x > 0 \quad \varphi(x) = De^{dx} + Fe^{-dx} \tag{11.60}$$

$$\text{III} \quad x \geqq a \quad \varphi(x) = Ce^{i\kappa x} \tag{11.61}$$

其中

$$\kappa = \sqrt{\frac{2mE}{\hbar^2}}, \qquad d = \sqrt{\frac{2m}{\hbar^2}(V-E)}$$

(11.59)式的 $Ae^{i\kappa x}$ 表示入射波,$Be^{-i\kappa x}$ 表示反射波。波函数 $\varphi(x)$是二阶微分方程 式的 Schrö-

dinger 方程式的解，所以 $\varphi(x)$ 和 $\varphi'(x)$ 必需是处处连续。即 $x=0$ 时，从 $\varphi(x)$，$\varphi'(x)$ 连续这一条件，分别得

$$A+B=D+F \tag{11.62}$$

$$i\kappa(A-B)=d(D-F) \tag{11.63}$$

又在 $x=a$，同样地由 $\varphi(x)$，$\varphi'(x)$ 连续这一条件可得

$$De^{da}+Fe^{-da}=Ce^{i\kappa a} \tag{11.64}$$

$$d(De^{da}-Fe^{-da})=i\kappa Ce^{i\kappa a} \tag{11.65}$$

从(11.62)～(11.65)式，可求得 B/A，D/A，F/A，C/A，其中 B/A 是入射波和反射波的振幅比，C/A 表示入射波和透射波的振幅比。且这些的平方各自对应于反射率、透射率。这里，以隧道效应穿过的粒子，最后穿透率 $|C/A|^2$ 由(11.62)～(11.65)式给出

$$\left|\frac{C}{A}\right|^2=\left\{1+\frac{V^2\sinh^2 da}{4E(V-E)}\right\}^{-1} \tag{11.66}$$

由(11.66)式可以看到，同 V 相比，E 愈小，另外势垒的厚度 a 愈大，穿透率则愈小，但决不是零，表示粒子穿过势垒转移的几率存在着有限值，现考虑 $E\ll V$ 的情形，则(11.66)式成为

$$\left|\frac{C}{A}\right|^2\cong e^{-2(a/\hbar)\sqrt{2m(V-E)}} \tag{11.67}$$

因而穿过的粒子数 N 为

$$N=N_0e^{-2(a/\hbar)\sqrt{2m(V-E)}} \tag{11.68}$$

这里、N_0 在区域 I 中，是每单位时间碰撞势垒的粒子数。

电子的隧道转移在生命过程中起着重要的作用。在 Gutmann[27] 研究了生物结构中电子隧道转移的可能机制后，Григоров 等[28] 在理论上研究了细胞色素，铁氧还原蛋白等电子传递络合链之间电子隧道转移的可能性。

Blumenfeld 等[29] 对于在生物体内发生电子移动，能否用通过隧道效应的电子转移来解释，作了半定量的分析。若势垒厚度 a（对应于电子传递体间的距离）为 1.0～2.0 nm，又把 $V-E$ 取为 1—2 eV，则得穿过的粒子数为 10^{-10}～10^{-2} 秒。必须指出，这样的计算毕竟是半定量的，能否说明问题以及能说明到何种程度等等都是需要进一步探讨的。

11.2.3 蛋白质在电子远距离转移过程中的作用

通过对一系列生物现象的研究已经了解到，当无载体参与时电子的迁移距离超过 3.0—7.0 nm。例如在细菌的光合作用中心，电子在光的供体和原始受体间的电荷分配过程中的迁移距离可达到 3.0—4.0 nm 的量级。电子在这样的距离上迁移是很难用简单的隧道机制来说明的。

从迄今为止的情况来看，电子的迁移过程往往是放在蛋白质系统中来加以研究的。为了说明电子迁移过程的分子机制，研究已知分子结构的分子间的迁移过程无疑是极其重要的。

研究表明，细胞色素 c 和亚铁血红蛋白间的氧化-还原反应是研究含铁蛋白质中电子迁移机制的方便模型。Атанасов 等[30] 对于电子在血红蛋白中的迁移过程进行了较详细的研究。以下，我们根据最简单的模型来讨论 α-螺旋蛋白质分子对供体和受体间的电子迁移的影响。

假定，由 $N(\gg 1)$ 个相距各为 a 的肽基组成的 α-螺旋蛋白质的电子状态由下列 Hamilton 算符描述：

$$H_0=\sum_n\left\{\varepsilon B_n^+B_n-\frac{1}{2}L(B_n^+B_{n+1}+B_{n+1}^+B_n)\right\} \tag{11.69}$$

式中 ε 是肽基中最弱的耦合电子能，n 是分子中肽基的号码，$\frac{1}{2}L$ 是相邻肽基的电子相互作用能。

设进入 n_1 肽基的供体分子能量算符为

$$H_D=\varepsilon_1 C_{n_1}^+C_{n_1},\ \varepsilon_1<\varepsilon \tag{11.70}$$

供体分子与肽基间的相互作用可写成

$$H_{int}=\delta_1(B_{n_1}^+C_{n_1}+B_{n_1}C_{n_1}^+) \tag{11.71}$$

算符 $B_n^+B_n$ 表征这样的状态，在该状态中电子处于 n 号肽基的能级 ε 上。$C_{n_1}^+C_{n_1}$ 表征这样的状态，在该状态中电子处于供体的能级 ε_1 上。算符 $B_n^+B_m$ 表示电子从肽基 m 向肽基 n 的跃迁，$C_{n_1}B_{n_1}^+$ 表示电子从供体向肽基的跃迁。

由于考虑到系统中的单电子状态，所以应有下列等式

$$C_{n_1}^+C_{n_1}+\sum_n B_n^+B_n=1 \tag{11.72}$$

系统的所有其它电子和原子核确定了我们的模型的链构型和参数 $\varepsilon,\varepsilon_1,\delta_1$ 和 L 的值。

因此，在我们的系统（供体＋蛋白质分子）中的最弱耦合作用的状态由下式 Hamilton 量确定

$$H=H_0+H_D+H_{int} \tag{11.73}$$

利用正则变换

$$\begin{aligned}B_n&=\frac{1}{\sqrt{N}}\sum_\kappa e^{-i\kappa na}A_\kappa\\ A_\kappa&=\frac{1}{\sqrt{N}}\sum_n e^{i\kappa na}B_n\end{aligned} \tag{11.74}$$

可将(11.73)写成

$$\begin{aligned}H=&\sum_\kappa[\varepsilon_\kappa A_\kappa^+A_\kappa+V_{n_1}(\kappa)C_{n_1}^+A_\kappa+V_{n_1}^*(\kappa)C_{n_1}A_\kappa^+]\\&+\varepsilon_1 C_{n_1}^+C_{n_1}\end{aligned} \tag{11.75}$$

式中 κ 是波函数在 $-\frac{\pi}{2a}<\kappa\leqslant\frac{\pi}{2a}$ 范围内取 N 个等间隔数值；ε_κ 是导带的 N 个能级的能量

$$\varepsilon_\kappa=\varepsilon-L\cos\kappa a \tag{11.76}$$

L 是导带的半宽度；

$$V_{n_1}(\kappa)=\frac{\delta_1}{\sqrt{N}}\exp(-i\kappa n_1 a) \tag{11.77}$$

可将系统的波函数写成

$$|\Psi_{n_1}\rangle=\left(u_1C^+_{n_1}+\sum_\kappa u_\kappa A^+_\kappa\right)|0\rangle \tag{11.78}$$

式中$|0\rangle$为无电子系统的状态，$A^+_\kappa|0\rangle$是电子处于导带的第κ能级时系统的状态，$C^+_{n_1}|0\rangle$是电子处于供体分子的能级ε_1时系统的状态。由(11.78)的归一化条件得等式

$$|u_1|^2+\sum_\kappa|u_\kappa|^2=1 \tag{11.79}$$

式(11.78)中的未知函数u_1和u_κ以及系统的能量E由泛函

$$J\equiv\langle\Psi_{n_1}|H|\Psi_{n_1}\rangle$$

的极小条件得到。将(11.78)及(11.75)代入上式可知，J的极小值可从解下列方程组得到

$$(\varepsilon_\kappa-E)u_\kappa+V^*_{n_1}(\kappa)u_1=0 \tag{11.80}$$

$$\sum_\kappa V_{n_1}(\kappa)u_\kappa+(\varepsilon_1-E)u_1=0$$

由方程组(11.80)的非平庸解的条件得等式

$$\varepsilon_1-E=\frac{\delta_1^2}{N}\sum_\kappa(\varepsilon_\kappa-E)^{-1} \tag{11.81}$$

它对应于系统的$N+1$个能级E。每一个这样的数值E对应的方程是

$$(E-\varepsilon_\kappa)u_\kappa(E)=V^*_{n_1}u_1(E) \tag{11.82}$$

由该式和(11.79)可以确定$u_\kappa(E)$和$u_1(E)$。特别是

$$u_1(E)=[1+\delta_1^2(E-\varepsilon_\kappa)^{-2}]^{-1} \tag{11.83}$$

当$\varepsilon-E>L$时，经过将(11.81)中对κ的求和换成积分后，可得

$$\varepsilon_1-E=\frac{\delta_1{}^2}{\sqrt{(E-\varepsilon)^2-L^2}} \tag{11.84}$$

在由(11.81)式确定的$N+1$个能级中，我们感兴趣的仅仅是最低能级E_1，其能量值与ε_1相差很小，可令$E_1=\varepsilon_1-\xi$，当$|\varepsilon-\varepsilon_1|>\xi$时，由式(11.84)得

$$\xi=\delta_1{}^2/\sqrt{(\varepsilon-\varepsilon_1)^2-L^2}$$

因而，与蛋白质分子相互作用的受体的电子能量数值为

$$E_1=\varepsilon_1-\delta_1{}^2/\sqrt{(\varepsilon-\varepsilon_1)^2-L^2} \tag{11.85}$$

根据(11.78)和(11.82)，与具有能量E_1的电子态相应的波函数具有形式

$$|\Psi^1_{n_1}\rangle=u_1(E_1)\left(C^+_{n_1}+\sum_\kappa\frac{V^*{}_{n1}A^+{}_\kappa}{E_1-\varepsilon_\kappa}\right)|0\rangle$$

将(11.77)中的$V_{n_1}(\kappa)$代入上式，并根据(11.74)变换算符$B^+{}_n$，就可写出在坐标表象中的$|\Psi^1{}_{n_1}\rangle$：

$$|\Psi^1{}_{n_1}(\boldsymbol{r})\rangle=u_1(E_1)\Big\{\varphi_D(\boldsymbol{r}-n_1\boldsymbol{a})$$
$$+\frac{\delta_1}{N}\sum_{\kappa,n}\frac{e^{i\kappa(n_1-n)a}\varphi_n(\boldsymbol{r}-n\boldsymbol{a})}{E_1-\varepsilon_\kappa}\Big\}$$

式中φ_D和φ_n分别是在自由供体分子和n号肽基中电子的波函数

$$\varphi_D(\boldsymbol{r}-n_1\boldsymbol{a})\equiv C^+_{n_1}|0\rangle,\ \varphi_n(\boldsymbol{r}-n\boldsymbol{a})\equiv B^+_n|0\rangle$$

将对 κ 的求和换成积分后波函数$|\varphi^1_{n_1}(\boldsymbol{r})\rangle$可写成

$$|\varphi^1_{n_1}\rangle=u_1(E_1)\left\{\varphi_D(\boldsymbol{r}-n_1\boldsymbol{a})+\delta_1\sum_n\frac{e^{-A_1|n_1-n|}\varphi_n(\boldsymbol{r}-n\boldsymbol{a})}{\sqrt{(\varepsilon-E_1)^2-L^2}}\right. \tag{11.86}$$

式中

$$A_1=\ln\frac{L}{\varepsilon-E_1-\sqrt{(\varepsilon-E_1)^2-L^2}} \tag{11.87}$$

这样，由于与蛋白质分子的相互作用，供体分子的波函数在整个蛋白质分子线度上不等于零，远距离处部分对函数的贡献随着$|n_1-n|$的增长以指数减小。在各种 E_1 值下，禁带宽度为 3.5eV 和能带半宽度为 0.5eV 的蛋白质的参数 A_1 的值列于表 11-2

表 11-2

E_1(eV)	3.00	3.20	3.40	3.45	3.48
A_1	3.84	2.78	1.85	1.56	1.32

现在假定，象与 n_1 肽基相连接的供体分子一样，具有振动激发能 ε_2 的受体分子也是与 n_2 肽基相连接的。这样，系统（供体、蛋白质分子、受体）的 Hamilton 量可以写成如下形式

$$H=\sum_\kappa\{\varepsilon_\kappa A^+_\kappa A_\kappa+\varepsilon_1C^+_{n_1}C_{n_1}+\varepsilon_2C^+_{n_2}C_{n_2}$$
$$+(V_{n_1}(\kappa)C^+_{n_1}A_\kappa+V^+_{n_2}(\kappa)C^+_{n_2}A_\kappa+\text{复共轭项})\} \tag{11.88}$$

式中

$$V_{n_2}(\kappa)=\frac{\delta}{\sqrt{N}}\exp(-i\kappa n_2a) \tag{11.88'}$$

定态波函数可写成如下形式

$$|\psi\rangle=(u_1C^+_{n_1}+u_2C^+_{n_2}+\sum_\kappa u_\kappa A^+_\kappa)|0\rangle \tag{11.89}$$

归一化条件为

$$|u_1|^2+|u_2|^2+\sum_\kappa|u_\kappa|^2=1 \tag{11.90}$$

未知函数 u_1，u_2 和 u_κ 以及能量 E 由下列齐次方程组确定

$$\begin{aligned}(\varepsilon_1+V_{11}-E)U_1+V_{12}U_2=0\\ V_{21}U_1+(\varepsilon_2+V_{22}-E)U_2=0\end{aligned} \tag{11.91}$$

式中

$$V_{11}(E)=\frac{\delta_1^2}{N}\sum_\kappa(\varepsilon_\kappa-E)^{-1},\ V_{22}(E)=\frac{\delta_2^2}{N}\sum_\kappa(\varepsilon_\kappa-E)^{-1}$$

$$V_{12}(E)=V_{21}(E)=\frac{\delta_1\delta_2}{N}\sum_{\kappa}\frac{\exp[-i\kappa(a|n_2-n_1|)]}{\varepsilon_\kappa-E}$$

当 $E<\varepsilon-L$ 时，将这些表示式对 κ 的求和换成积分，可得

$$V_{11}(E)=\frac{\delta_1{}^2}{\sqrt{(\varepsilon-E)^2-L^2}}$$

$$V_{22}(E)=\frac{\delta_2{}^2}{\sqrt{(\varepsilon-E)^2-L^2}} \tag{11.92}$$

$$V_{12}(E)=V_{21}(E)=\frac{\delta_1\delta_2}{\sqrt{(\varepsilon-E)^2-L^2}}\exp\{-\Lambda(E)|n_1-n_2|\} \tag{11.93}$$

式中

$$\Lambda(E)=\ln\frac{L}{\varepsilon-E-\sqrt{(\varepsilon-E)^2-L^2}} \tag{11.94}$$

由方程组（11.91）的非平庸解的条件得到确定 $N+2$ 个能级系统的方程为

$$[\varepsilon_1+V_{11}(E)-E][\varepsilon_2+V_{22}(E)-E]=V_{12}{}^2(E) \tag{11.95}$$

方程$[\varepsilon_1+V_{11}(E)-E]=0$和$[\varepsilon_2+V_{22}(E)-E]=0$确定包含与蛋白质分子相互作用的供体或者受体分子系统的能量状态。式(11.95)中的函数 $V_{12}(E)$在能量为 E 的状态中表示了通过蛋白质分子实现供体和受体间的有效相互作用。

为简化起见，我们假定自由受体分子的基态电子能量 ε_1 等于$\hbar\Omega$，以使其第一振动激发与供体基态能量一致，即 $\varepsilon_2=\varepsilon_1$。再假定 $\delta_1=\delta_2$，这时 $V_{11}(E)=V_{22}(E)$，由(11.95)得到确定系统定态的方程式

$$E=\varepsilon_1+V_{11}(E)+V_{22}(E) \tag{11.96}$$

靠近能量 ε_1 的定态能量由下式确定

$$E_{\pm}^{(1)}=\hbar(\omega_0\pm\frac{1}{2}f) \tag{11.97}$$

其中

$$\hbar\omega_0\equiv V_{11}(\varepsilon_1),\quad \frac{1}{2}\hbar f\equiv V_{12}(\varepsilon_1) \tag{11.98}$$

这些状态的波函数可通过供体和受体分子的(11.86)类型的波函数表示出来

$$|\Psi_\pm(t)\rangle=\frac{1}{\sqrt{2}}\ [\psi^{D}{}_{n_1}\pm\psi^{A}{}_{n_2}]\ \exp\{-i(\omega_0\pm\frac{1}{2}f)t\} \tag{11.99}$$

在定态(11.99)中，电子以相同的几率分布于供体的基态与受体的振动态之间。

构造函数

$$\begin{aligned}\Psi&=\frac{1}{\sqrt{2}}(|\psi_+(t)\rangle+|\psi_-(t)\rangle)\\&=[\psi^{D}{}_{n_1}\cos ft/2-i\psi^{A}{}_{n_2}\sin ft/2]\end{aligned} \tag{11.100}$$

当 $t=0$ 时，这个函数描写电子处于供体中时系统的状态。在时刻 $t=\pi/f$，电子转移至受体分子。然而，当 $t=2\pi/f$ 时，又重新回到供体。因此电子以频率 f 来回于供体和受体分子之间。

只有当频率 $\nu>f$ 的受体分子的振动激发能 $\hbar\Omega$ 转入耗散系统中去时，才可能发生电子从供体向受体的真实转移。

当存在耗散介质时，电子从供体分子向受体分子转移的研究可完全象中性激发转移过程的研究一样地进行。

设 $t=0$ 时，电子处于供体的电子基态(11.86)，根据(11.100)式，f 值的确定不考虑耗散时的供体和受体间的激发交换频率。设 ν 是受体中振动激发消失的耗散过程的频率。当$\nu>f$ 时，按照文献[31]，电子进入受体振动态的几率由下式确定

$$w_{\mathrm{v}}(t)=\frac{f^2}{\nu^2-f^2}e^{-\nu t}\left[sh\left(\frac{t}{2}\sqrt{\nu^2-f^2}\right)\right]^2$$

$$\approx\frac{f^2}{\nu^2-f^2}\left[e^{-\frac{f^2 t}{2\nu}}-2e^{-\nu t}+e^{-2\nu t}\right] \tag{11.101}$$

电子进入受体电子基态的几率随时间的变化为

$$w_{\mathrm{G}}(t)=1-e^{-\nu t}\left\{1+\frac{\nu^2}{f^2-\nu^2}\left[1-ch(\sqrt{\nu^2-f^2}t)\right]\right.$$

$$\left.+\frac{\nu}{\sqrt{\nu^2-f^2}}sh(\sqrt{\nu^2-f^2}t)\right\} \tag{11.102}$$

因此，当 $t\to\infty$，$w_{\mathrm{v}}(\infty)=0$，$w_{\mathrm{G}}(\infty)=1$。

如果 $f>\nu$，则相应的几率由以下函数确定

$$w_{\mathrm{v}}(t)=\frac{f^2}{f^2-\nu^2}\sin(\sqrt{f^2-\nu^2}t/2) \tag{11.103}$$

$$w_{\mathrm{G}}(t)=1-e^{-\nu t}\left\{1+\frac{\nu^2}{f^2-\nu^2}[1-\cos(\sqrt{f^2-\nu^2}t)]\right.$$

$$\left.+\frac{\nu}{\sqrt{f^2-\nu^2}}\sin(\sqrt{f^2-\nu^2}t)\right\} \tag{11.104}$$

尤其是，当不存在弛豫($\nu\approx0$)时

$$w_{\mathrm{v}}(t)=\sin^2(ft/2),\ w_{\mathrm{G}}(t)=0$$

由以上讨论可以看到，蛋白质分子是可以促进电子从供体向受体分子转移的。但是，必须指出：(1) 对于蛋白质分子的未填满的电子状态，在理论中是仅仅作为虚拟状态被考虑的。电子进入受体分子，而不进入导带。(2) 蛋白质分子的影响仅仅表现为，供体和受体的波函数被改变了。在自由供体与受体分子中电子状态的波函数基本上定域于分子体积中。如果它们被固着于准周期结构的蛋白质分子上，则波函数具有“延伸的尾巴”[见式(11.86)]。这些函数的空间重叠保证着电子的转移。

以上我们介绍了 Лавыдов 等对蛋白质在电子远距离转移过程中的作用所作的某些研究。虽然在这一节里，我们没有对他们的工作做出专门的评价，但若干研究结果说明了他们的

研究是比较严格的，有意义的。

11.2.4 质子转移与氢键

无数的研究表明，生物体中的许多基本过程不仅与电子的转移相联系，而且也与质子的转移有关。Kirkwood 指出，生物高分子如蛋白质和核酸等，分子表面上有大量的质子。这些质子不断地运动着，在质子转移的同时也就伴随着激发能的转移。生物大分子与周围分子的相互作用，也可能受到质子运动的影响。质子运动可以与邻近类似的分子相呼应，从而构成能量转化和传递的系统，使诸如 DNA 这样的分子能够利用这种传能方式来传递信息[11]。

然而，从已经进行的研究工作来看，我们感兴趣的是通过氢键的质子转移，以及这种转移与能量转移的联系，直至这种转移对生物大分子结构和功能的影响等问题。

通过对蛋白质和核酸结构与功能的生物化学、生物物理学等的研究，我们已经在一定程度上了解到氢键的重要性(参见 3.2.2)。

现以 DNA 分子为例。我们知道，DNA 分子的两条单链以氢键维系着而形成严格有序的双螺旋结构。研究表明，氢键中质子的摆动也能转移能量，它摆动的能量子波长相当于红外线 2.7 μm 左右，与三线态和高能磷酸键的能量都十分一致。在 DNA 复制和转录过程中，作为反应原料的三磷酸核苷中高能磷酸键的能量就很容易转移到氢键上而与模板链上的核苷酸进行配对，形成新的 DNA 或 mRNA 分子。此外，氢键两端的 π 电子云有足够的重迭，π 电子可以横越氢桥而实现非定域化，有助于实现激发能的共振转移和半导性能带的电子转移。

研究表明，DNA 分子中的氢键形成基团与水的四配价晶格恰好相符，因而可以产生特殊的偶合，使水分子按一定有序的取向排列。其“长程有序化效应”可达 100nm 的范围，有助于形成 DNA 分子周围的液晶态有序水结构，这对形成三线态和激子转移等均十分有利。任何外界影响，如果破坏这种体系中局部水的平衡结构，就必然会通过结构重排列而传递至远距离，传递的机制即按质子转移的方式进行。

Lowdin 曾提出 DNA 分子中构成氢键的质子具有隧道效应（有关质子隧道效应以及质子隧道效应引起基因突变的假说请参见第 7 章）。

在氢键系统中，假如发生质子转移，随之发生电子转移，是理所当然的，然而电子转移不那么大。可是，在生物体内和质子转移偶合引起电子转移，帮助了某些反应的进行的例子是有的，例如以水为媒介的质子转移和糜蛋白酶活性中心通过电荷转接系统的质子转移与电子转移的偶合等等。

由于质子转移在一些基本化学过程（如酸-碱平衡、含水离子传导和亲核加成反应等）和生物学过程（如光合作用、膜上的输送、ATP 的形成和酶催化等）[32-37]中起着重要的作用，因而有关伴随质子转移而发生的能量转移的研究受到了愈来愈多的科学工作者的关注。譬如，近些年来，Scheiner 等人对氢键系统中的质子转移所作的研究即是这些研究中的一部分[38-42]。

Scheiner 运用 ab initio 分子轨道法对小模型系统$(H_3NHNH_3)^+$(参见蛋白质和核酸的结构）作了较深入的理论研究。得到的电子密度差异图清楚地表明，电子的再分配伴随质子从其平衡位置(NH—N)向氢键(N—H—N)的中点的半转移(参见图 11-5 至图 11-7)。



图 11-5 在$(H_3NHNH_3)^+$中，中央质子半-转移示意图

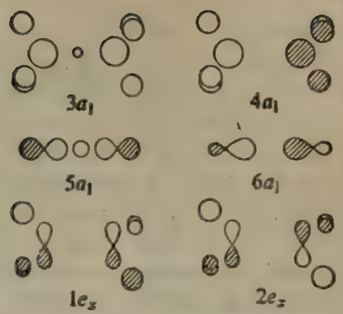


图 11-6 分子轨道示意图



图 11-7 分子轨道($5a_1$，$6a_1$)对的电子密度差异图

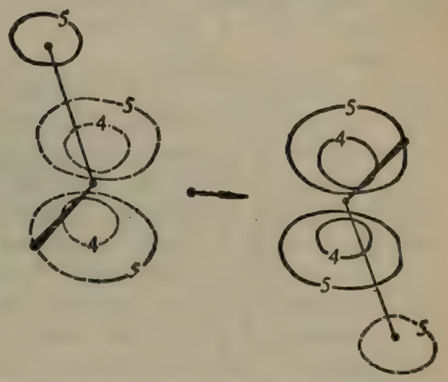


图 11-8 ($1e_x$，$2e_x$)对的密度差异图

从不同分子轨道总密度分布的改变可以看出，a_1 轨道跟沿 H-键轴线的密度改变相关（图 11-7 即是电子密度的改变跟分子轨道($5a_1$，$6a_1$)对的相关模型）。e 对称轨道引起 NH 键的极化且密度改变正交于 H-键轴线。

须要提及的是，1978年诺贝尔奖获得者 Mitchell 指出[43]，在生物体的氧化磷酸化过程中，由于呼吸链酶系的作用，使膜的两侧形成质子梯度差，这个梯度差就是生物能量的来源。

总之，“生物系统中能量的转移”作为量子生物学中的一个研究领域已经取得了一些有意义的结果，它将会引起更多生物学家、物理学家和化学家等的兴趣。他们将抓住能量转移的理论和实验研究，以帮助他们探索生物分子的结构及其相互作用，进而探索生命的奥秘。

参考文献

[1] A. 圣-乔其著，林克椿译：《生物能力学》，科学出版社，(1962 年).

[2] 徐京华：科学通报，(4):114(1960).

[3] Давыдов, А. С.: “Биология и Квантовая Механика”, (Наук. Думка, Киев), (1979).

[4] 程极济译：《蛋白质及核酸的激发态与能量转移译文集》，科学出版社，(1981).

[5] 程极济等：《生物物理学》，人民教育出版社，(1981).

[6] Förster, Th.: Ann. Phys., 2:55 (1984).

[7] Förster, Th.: Disc. Faraday Sec., 27:7 (1959).

[8] Stryer, L. et al.: Proc. Natl. Acad. Sci. U. S. A., 58:719 (1967).

[9] Bücher, H. et al.: Mol. Cryst., 2:199 (1967).

[10] Bennett, R. G.: J. Chem. Phys., 41:3037 (1964).

[11] 王身立：«控制论与遗传进化问题»，湖南科学技术出版社，(1985).
[12] Frenkel, J. I.: Phys. Rev., 37:19 (1931).
[13] Агранович, В. М.: "Теория Экситонов", (Наука, Москва), (1968).
[14] Moffitt, W.: Proc. Natl. Acad. Sci. U. S. A., 42:736 (1956).
[15] Давыдов, А. С. И Др.: Укр, Физ. Журн., 22:881 (1977).
[16] Davydov, A. S. et al.: Phys. Status Solidi (b), 59:465 (1973).
[17] 庞小峰：生物化学与生物物理学报，18:1 (1986).
[18] Scott, A. C.: Phys. Lett., 86(A):60 (1981).
[19] Scott, A. C.: Phys. Scripa., 25:651 (1982).
[20] 吴祖跃：大自然探索，4:33 (1985).
[21] 永田亲义著．陶宗晋、江寿平译：«量子生物学入门»，上海科学技术出版社，(1979).
[22] A. 圣-乔其著．林克椿译：«亚分子生物学导论»，科学出版社，(1966).
[23] 严继民：化学通报，(11):1 (1983).
[24] Mulliken, R. S. et al.: "Molecular Complexes: A Lecture and Reprint Volume", (John Wiley, New York), (1969).
[25] Guryanova, E. N. et al.: "Donor-Acceptor Bond", (John Wiley, New York), (1975).
[26] Flurry, R, L.: J. Phys. Chem., 69:1927 (1965).
[27] Gutmann, F.: Nature, 219:1359 (1968).
[28] Григоров, Л. Н. И Др.: Биофизика, 17:195 (1972).
[29] Blumanfeld, L. A. et al.: J. Theor. Biol., 39:1 (1973).
[30] Атанасов, Б. Л. И Др.: Молекуляр. Биология., 11:537 (1977).
[31] Давыдов, А. С.: "Квантовая. Механика", (Наука, Москова), (1973).
[32] Stoeckenius, W.: Sci. Am., 234:38 (1976).
[33] Henderson, R.: Ann. Rev. Biophys. Bioeng., 6:87 (1977).
[34] Sauer, K.: Acc. Chem. Res., 11:257 (1978).
[35] Mitchell, P.: Science, 206:1148 (1979).
[36] Cohen, S. M. et al.: Nature, 273:554 (1978).
[37] Jenck, W. P.: "Catalysis in Chemistry and Enzymology", (McGraw-Hill, New York), (1969).
[38] Scheiner, S.: J. Am. Chem. Soc., 103:315 (1981).
[39] Scheiner, S.: Ann. NY. Acad. Sci., 367:493 (1981).
[40] Scheiner, S. et al.: J. Am. Chem. Soc., 103:2169 (1981).
[41] Scheiner, S. et al.: Chem. Phys. Lett., 79:39 (1981).
[42] Scheiner, S.: Int. J. Quant. Chem., QBS., (8):221 (1981).
[43] Mitchell, P.: FEBS Lett., 78:1 (1977).

沈仲钧
刘次全
王身立
王　莹

第12章　化学致癌作用的量子生物学研究

12.1 概　　述

12.1.1 历　　史

在量子生物学中，有关化学致癌剂的致癌机理的研究占据了十分重要的地位，亦取得了不少有意义的结果，因而愈来愈受到科学家们的普遍重视。

早在 1937 年 Schmidt 就提出了致癌物分子的电子结构与其致癌活性有关的设想。他用箱型模型法近似计算了芳香族碳氢化合物中 π 电子的分布。结果显示，这些分子中，中位 π-电子分布特别大的，其致癌作用亦强。Schmidt 进而认为这个位置与致癌有关[1]（图 12-1）。

进入 40 年代后，Pullman 夫妇等[2]又提出了芳香族碳氢化合物致癌的 K-区理论。他们用定域的 Hückel 分子轨道法比较了芳香族碳氢化合物的致癌作用与反应指数的关系。结果看到，重要的不是中位，而是所谓的"K-区域"（图 12-2）。



图 12-1　苯并[a]蒽的"中位"

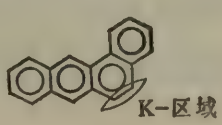


图 12-2　苯并[a]蒽的"K-区域"

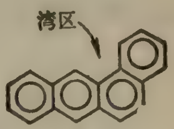


图 12-3　苯并[a]蒽的"湾区"

70 年代，Jerina 等提出了化学致癌的"湾区理论"[3-5]。该理论是在生物实验的基础上提出的，它用湾区碳阳离子的稳定性作为判据，较好地解释了多环芳烃（以下简称 PAH）最终致癌形式的代谢历程（参见图 12-3）。

1979 年，我国科学工作者戴乾圜提出了以离域能为参数的"双区理论"，建立了计算芳香族碳氢化合物致癌活性的定量公式[6]。

以上提到的几种理论，用于解释 PAH 的结构与致癌活性关系方面，取得了不同程度的成功，但致癌物是多种多样的，致癌过程是受多种因素影响的，因而对于化学致癌剂的致癌机理的研究，也先后提出了其它的一些理论，如西德量子化学家 Ladik 提出的"误读理论"等等。但是，要真正弄清致癌这样一个复杂的问题，并非是某一个单一理论所能及的，为探索致癌之谜，各种不同学说的提出和各种不同方法的运用是很自然的，研究实践说明了这一点。

12.1.2 化学致癌物的发现

二百多年来，随着工业，特别是化学工业的发展，人们逐渐发现了一些化学物质对人或动物的致癌作用。

早在 1775 年，英国外科医生 Pott[7]发表了他对伦敦扫烟囱工人的阴囊癌的观察，指出了这种阴囊癌是烟囱里的烟灰长期积沉在阴囊皱襞中的结果。从这以后，人们开始注意某些肿瘤与职业的关系。1895 年 Rehn 报告了接触苯胺染料的工人易患膀胱癌，这个事实曾引起了当时社会上激烈的争论。后经研究，证实了某些化学物质与癌症病因有密切关系。1915 年，日本人山极和市川[8]用煤焦油涂抹兔耳及小白鼠的皮肤，成功地引发了皮肤癌，几年以后，这个现象又进一步得到了证实和发展。1918 年 Tsutsui[9]及 1922 年 Passey[10]用烟灰的有机溶剂提取液涂抹小鼠皮肤致癌成功又进一步证实了 Pott 的观察。1932～1933 年，Kennaway 和 Cook[11]等人，在多年研究的基础上先后证明二苯并蒽、苯并[a]芘等 PAH 是煤焦油中致癌的化学成分。随着煤焦油化学工业的发展，特别是芳香烃和偶氮染料工业的发展，各国科学家采取实验肿瘤的方法对芳香烃类、偶氮类、亚硝基化合物进行了大量的动物致癌实验。到 1970 年已对五千多种化合物进行了测试，与此同时，还进行了肿瘤死亡统计，直到 1978 年已肯定了近三千种对动物有致癌作用的化合物，而且以后每年都有数以百计的新致癌物被报道。这些物质，相当多的种类存在于生物圈里，直接威胁着人类的健康。

迄今为止，就已经发现的致癌物而言，有无机的，也有有机的；有结构简单的，也有结构复杂的；有直链的，也有环状的；有人工合成的，也有天然产品等等。归纳起来可以分为：

多环芳烃类，芳胺类，偶氮类，卤烃类，亚硝胺及亚硝酰胺类，烷化剂，尿烷及硫脲类，内脂类，性激素类，霉菌毒素，植物产品，无机化合物及其它致癌性物质。

12.1.3 体内、体外代谢实验与最终致癌物

12.1.3.1 代谢途径

体内和体外代谢实验表明，化学致癌物一般可分为直接致癌物和间接致癌物两大类。

1. 直接致癌物

直接接触机体某一部位能引起癌瘤的化学物质一般称为直接致癌物。在人类环境中，这类致癌物为数极少。

大多数直接致癌物为化学合成产物，它们可能构成一定的职业性危险。由于它们在化学上较活泼，故一般在环境中维持时间不长，但它们在体内能释出亲电子的反应物，同细胞的大分子受体结合。如图 12-4 所示，为氮芥同鸟嘌呤 N7 位置的结合。

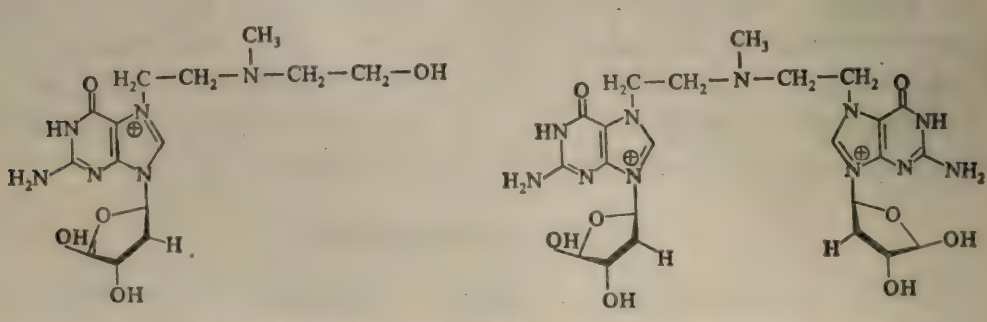


图 12-4 氮芥同鸟嘌呤 N7 位的结合示意图（上为单加成物，下为交联物）

2. 间接致癌物

需要通过机体代谢活化才能引起瘤瘤的化学物质一般称为间接致癌物。

大多数致癌物为间接致癌物。它们在环境中相对是稳定的，进入机体后才代谢活化，所以污染环境的危险较多。对间接致癌物代谢的研究，是从生化药理学得到启发的。它们在体内的代谢具有双重性，一方面可经代谢活化为最终致癌物，另一方面可代谢而失去致癌性。图 12-5 为化学致癌物代谢致癌示意图。

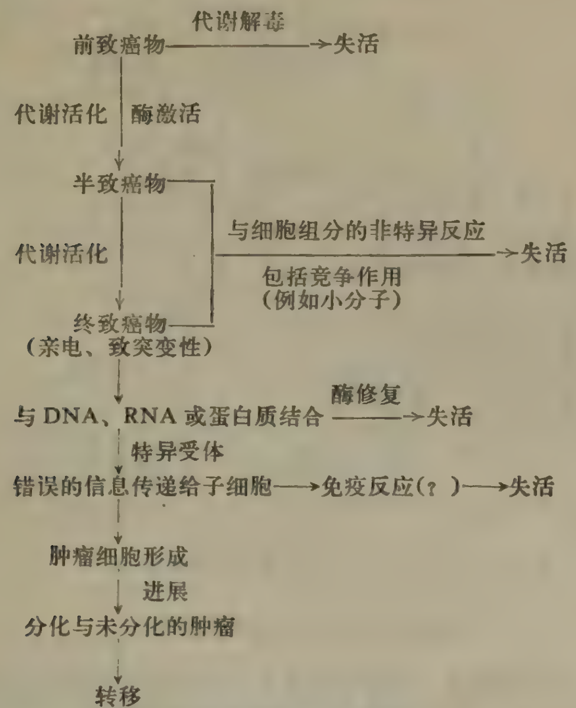


图 12-5 化学致癌物致癌示意图

间接致癌物主要包括多环芳烃和杂环芳烃、黄曲霉毒素、芳香胺、亚硝胺、偶氮染料和卤代烃等等。其中以多环芳烃为最重要的环境致癌物。它的种类之多，分布之广，与人类关系之密切是任何其它环境致癌物所不能比拟的。迄今为止，量子生物学对化学致癌的研究也主要集中在 PAH 上。因此，在这一章里我们将着重以 PAH 为例来予以讨论。

图 12-6 表示苯并[a]芘*可能的代谢途径。

12.1.3.2 最终致癌物

从图 12-6 和 12-7 可以看到，苯并[a]芘的代谢途径和在机体内的活化过程。

迄今，若干实验证据有力地暗示，湾区四氢二醇环氧化物（7,8,9,10-四氢-7,8-二羟基-9,10-环氧化苯并[a]芘）系苯并[a]芘的半致癌形式，即它所产生的一特别重要的转化——通过质子化和环氧环的打开而形成三羟基碳阳离子，即苯并[a]芘的最终致癌物（图 12-7）。

* 是 PAH 中研究得最为详尽的一种强烈致癌剂。

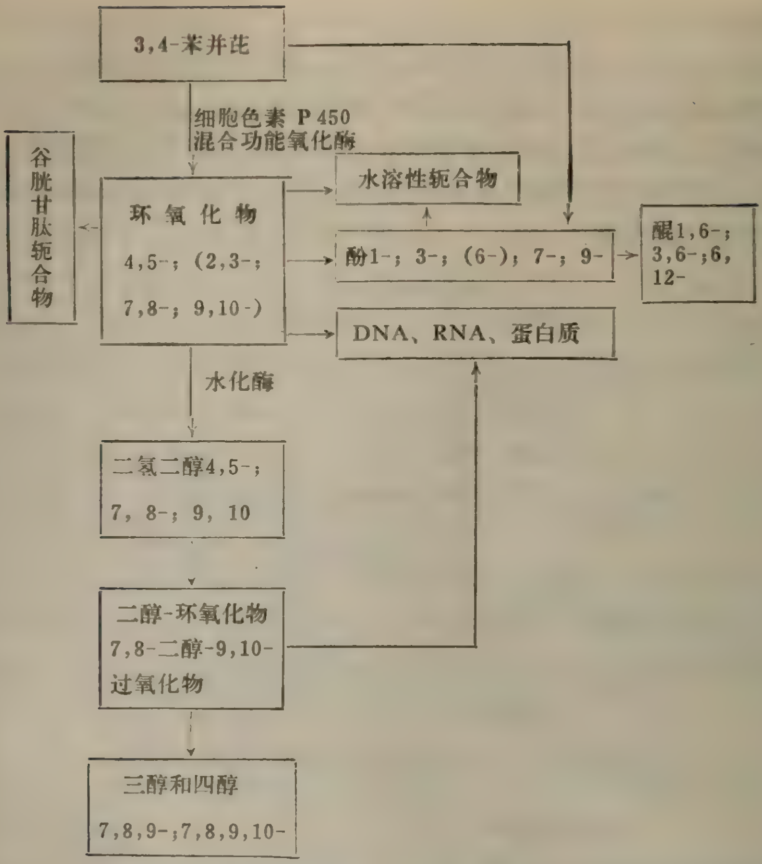


图 12-6 苯并[a]芘可能的代谢途径

概括起来，化学致癌作用的量子生物学研究自 Schmidt 提出分子内的电子分布与致癌关系的设想以来，到现在已经取得了不少成果，但仍在继续深入之中。目前，肿瘤已成为医学、分子生物学、量子生物学、细胞生物学、免疫学、遗传学等多种学科的研究对象。当今世界正汇集各方面的科学家，在电子、分子、细胞、个体以及群体水平，从流行病学、病因学、临床诊断和治疗，以及包括致癌机理、生物学特性和基础理论方面开展广泛深入的研究。

12.2 结构与活性关系研究

12.2.1 多环芳烃(PAH)

在 12.1.1 中，我们提到了化学致癌作用研究中的“K-区理论”，“湾区理论”，“双区理论”等等，在讨论 PAH 的这一节里，我们将比较详细地介绍这些理论。

12.2.1.1 L-区理论

在 30 年代末，Schmidt 提出了致癌物分子的电子结构与其致癌活性有关的设想。他们认为，

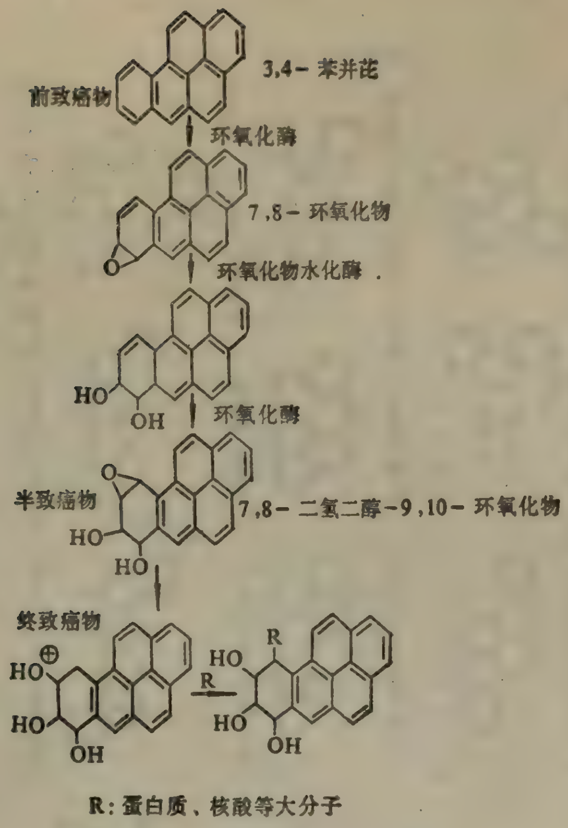


图 12-7 苯并[a]芘在机体内的活化过程

癌变是由于细胞受到某些外来因素的催化而处于某种激发状态[1]，并用价键法计算了芳香族碳氢化合物中 π 电子的分布,进而以 π 电子的自由价作为结构指数，通过比较，发现在这些分子中,中位（相当于蒽和类似分子的 9、10 位置）π 电子密度分布特别大的，其致癌作用亦强。他们进而认为,这个位置(即 L 区)与致癌过程有关(参见图 12-1)。但是，研究表明(图 12-8),中菲位电子密度大的 PAH 可以在蛋白质的酮-烯醇互变异构过程中起一种催化作用（参见 12.2.1.2)；这样就导致了细胞内蛋白质分子的不可逆改变,从而促进细胞癌变。

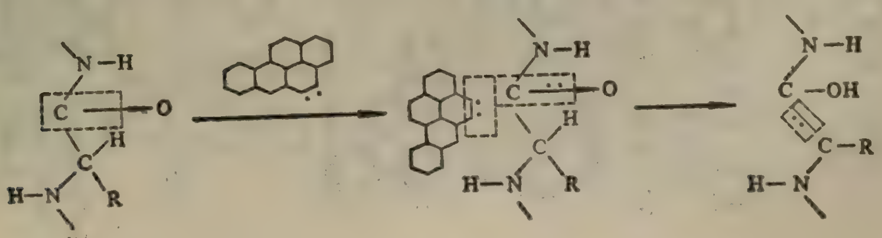


图 12-8

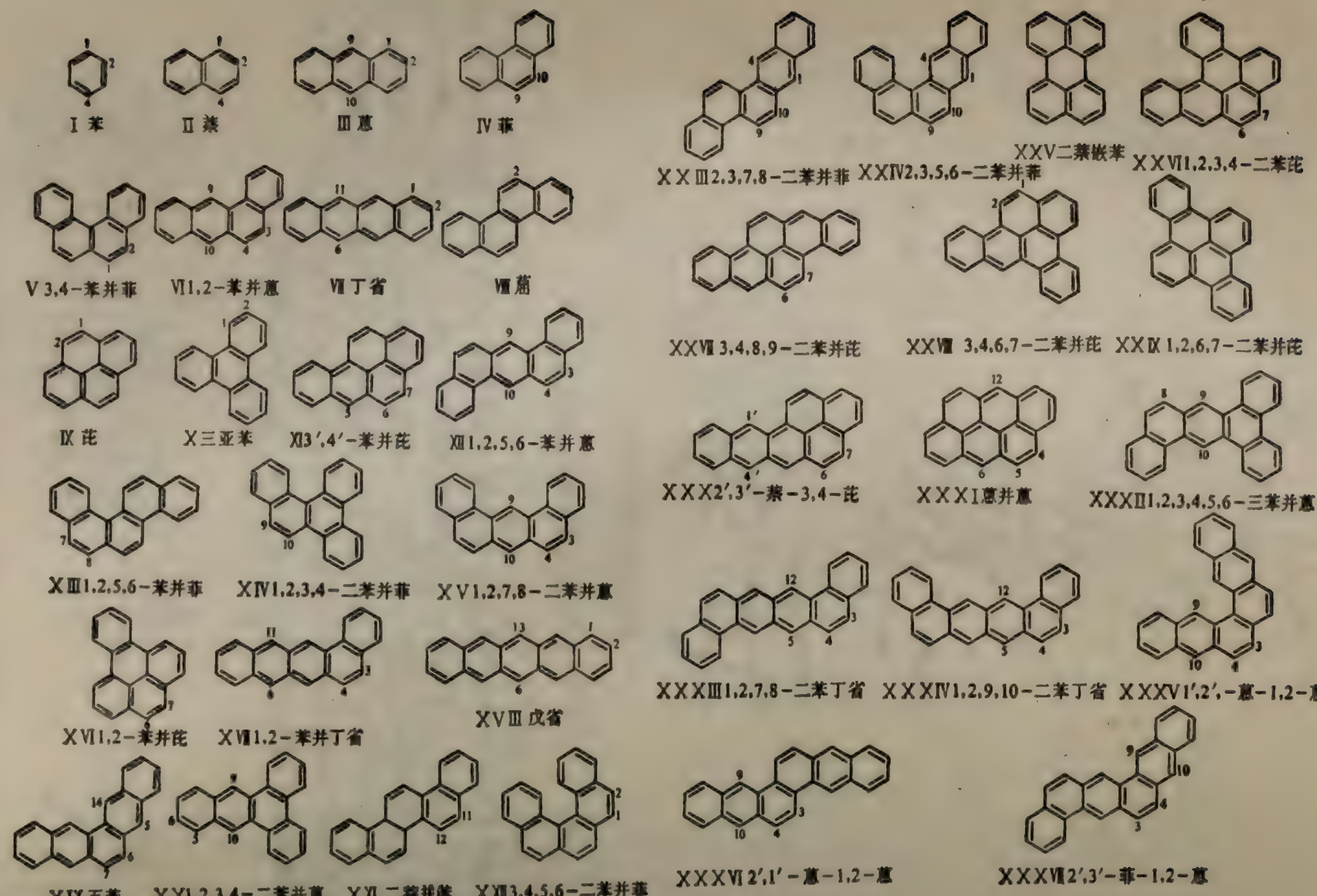


图 12-9 未取代芳香族碳氢化合物

表 12-1　未取代芳香族碳氢化合物的致癌作用和

K-区域，L-区域的理论指数

化合物	致癌性	K-区域		L-区域	
		键	O.L.E.+C.L.E.$_{min}$	位置	P.L.E.+C.L.E.$_{min}$
I	—	1-2	4.07	1-4	6.54
II	—	1-2	3.56	1-4	5.98
III	—	1-2	3.53	9-10	5.38
IV	—	9-10	3.36	—	—
V	+	1-2	3.41	—	—
VI	+	3-4	3.29	9-10	5.53
VII	—	1-2	3.33	6-11	5.25
VIII	—	1-2	3.38	—	—
IX	—	1-2	3.33	—	—
X	—	1-2	3.81	—	—
XI	++++	6-7	3.23	—	—
XII	++	3-4	3.30	9-10	5.69
XIII	+	7-8	3.41	—	—
XIV	+	9-10	3.42	—	—
XV	+	3-4	3.31	9-10	5.66
XVI	—	6-7	3.37	—	—
XVII	—	3-4	3.20	6-11	5.25
XVIII	—	1-2	3.27	6-13	5.03
XIX	—	6-7	3.23	5-14	5.56
XX	—	5-6	3.51	9-10	5.67
XXI	—	11-12	3.37	—	—
XXII	—	1-2	3.38	—	—
XXIII	—	9-10	3.27	1-4	5.47
XXIV	—	9-10	3.30	1-4	5.48
XXV	—	—	—	—	—
XXVI	++++	6-7	3.24	—	—
XXVII	++++	6-7	3.17	—	—
XXVIII	+	1-2	3.34	—	—
XXIX	—	—	—	—	—
XXX	—	6-7	3.14	1′-4′	5.30
XXXI	—	4-5	3.20	—	—
XXXII	—	7-8	3.35	9-10	5.80
XXXIII	—	3-4	3.24	5-12	5.42
XXXIV	—	3-4	3.25	5-12	5.40
XXXV	—	3-4	3.30	9-10	5.45
XXXVI	—	3-4	3.24	9-10	5.44
XXXVII	—	3-4	3.24	9-10	5.54

12.2.1.2 K-区理论

在 40 年代末和五十年代初，Pullman 等[2]在 Schmidt 工作的基础上，提出了开创性的 K-区理论。他们运用定域的 HMO 法比较了 37 个未取代芳香族碳氢化合物的邻位定域能（O.L.E.），对位定域能(P.L.E.)，碳定域能(C.L.E.)，（见图 12-9 、表 12-1 ），结果看到，重要的不是 L 区，而是所谓的“K 区”。概括起来，K 区理论最重要的两点结论是：

(1) K区是发生致癌反应的关键区域，而L区是对致癌反应起抗拮作用的区域；

(2) 致癌反应的本质是致癌剂和生物体中各种亲电子活性中心发生加成反应，此时电子的流向是从致癌剂⟶生物体成分。

Pullman 等在对K区和L区结构指数计算的基础上，分析了它们与致癌活性的关系。起初十分失望，没有见到一种简单的指标能够有效地反映与致癌活性的联系。后来，他们选择了一种复合指标来标志反应能力。对K区用键定域能*（扰乱一个共轭分子的电子结构，使一对 π 电子定位在两个相邻碳原子间所需的能量）和碳定域能(使两个电子定位在发生反应的一个碳原子上因而不再参与共轭所需的能量）之和；对L区用对位定域能(使两个电子同时定位在碳原子所需的能量）和碳定域能之和。在经过计算和比较之后，他们提出了定量的判据：凡K区的键定域能与碳定域能之和（K区复合指标）小于或等于 3.31 β，L区的对位定域能与碳定域能之和（L区复合指标）大于或等于 5.66 β 时，分子才可能有致癌活性（参见表 12-1)。

从表 12-1 可以看出，3,4-苯并芘（苯并[a]芘)无L区，K区复合指标为 3.23 β，是很强的致癌物；1,2,5,6-二苯并蒽K区复合指标为 3.30 β，而L区为 5.69 β，只略大于 5.66 β，为次强的致癌物；1,2-苯并蒽（苯并[a]蒽）K区复合指标是 3.29 β，L区复合指标为 5.53 β，前者小于 3.31 β，后者小于 5.66 β，故仅有微弱的致癌作用。总之，按照K区理论的计算，与当时已知的 37 个 PAH 的实验事实较为一致。

由于K区理论大致解释了当时已知的不多的事实，在一定意义上开辟了将分子轨道理论用于生理活性研究上的可能途径，一度曾被认为是结构与生理活性关系问题上的一项重大突破。但随后的实验表明，K区理论的许多预言与实验事实是相矛盾的，且根据K区理论提供的线索，合成出来的K区环氧化物，总是致癌活性较低。因而，使人们对K区理论产生了怀疑，渐渐失去了影响。

分析K区理论未能取得成功的原因，其主要不足在于缺乏 PAH 在生物体内实际代谢过程的充分资料。这也说明理论计算必须根据实验结果加以修正和发展；另一方面，理论计算也为实验工作者提供了在某种程度上的指导性线索、预测和计算工具。

12.2.1.3 湾区理论

60 年代期间，许多科学工作者，尤其是 Lacassagne 及其同事进行了 PAH 的结构致癌活性关系的实验研究[12]。七十年代以来，致癌剂尤其是 PAH 生物代谢实验资料的大量积累[13-17]，使人们认识到角环的环氧化物是与生物大分子共价结合的重要中间体。进一步的工作表明，处于所谓湾区(Bayregion)的角环(如图 12-10 所示)在代谢活化过程中对致癌性能起着关键作用，

---

* 键定域能即邻位定域能。

而其最终致癌形式，是湾区环氧化物(参见图 12-7)。在这些实验事实的基础上，Jerina 提出了“湾区理论”[3-5]，其主要论点是：

(1) 湾区的角环在代谢活化过程中，对致癌反应起着关键性的作用，其最终致癌形式为湾区环氧化物。

(2) 湾区碳阳离子稳定性越高，致癌性越强。



图 12-10 苯并[a]芘的湾区

湾区理论强调湾区碳阳离子的重要性(图 12-11)。Jerina 等认为，PAH 湾区碳阳离子的离域能越高，其致癌活性越强。已知某些湾区环氧化物的二氢二醇前体，其致突变或致癌活性与母体 PAH 相接近，因此湾区理论受到了人们的重视。

湾区碳阳离子和其它位置阳离子的稳定性，可以借助微扰分子轨道法 (PMO 法) (参见 2.2.1.2)计算阳离子的离域能 ($\Delta E_{离域}$)而定量地估计。理论计算表明，在 Jerina 等所选用的 25 个 PAH 中，除个别例外，湾区碳阳离子的离域能与实验致癌标度尚属平行。但对更多的PAH作理论计算后，则发现有很多例外，如一些有多个湾区的化合物没有致癌活性 (图 12-9 和表 12-1 中的 1,2,6,7-二苯并芘等)，而另一些没有湾区的化合物则有程度不等的致癌活性。由此可见，只是形式地从分子中存在湾区与否的角度来确定是否致癌的问题，并未抓住问题的实质。况且，“湾区理论”也没有提出任何定量判据，因而也缺乏预言能力[6]。

12.2.1.4 双区理论

1979 年以来，我国学者戴乾圜对 PAH 致癌机理进行了较系统的分子轨道计算和其它一些研究[6,12,19]，提出了以离域能为参数的“双区理论”，建立了计算 PAH 致癌活性的定量公式。其主要结论是：

(1) PAH 分子致癌的必要结构是存在着两个且仅需两个活性区域，即致癌剂在代谢过程中将产生两个亲电活性中心。换句话说，就是致癌剂在体内分步地或同步地代谢成双官能烷化剂是致癌的最有效步骤。

活性区域包括：角环、次角环和活性K区(图 12-12)。

(2) 脱毒与活化是两个相关而又相互矛盾的过程。实际上，角环、K区，甚至L区都有活化和脱毒的双重作用。其中，角环、次角环、活性K区，主要起代谢活化的作用，而L区、等价K区及超过双区需要以上的多余的角环，主要起脱毒作用。K区是一种双重性能区域(参见12.2.1.8)。

(3) 致癌的关键过程是要求 DNA 股间互补碱基对的横向交联，以及致癌的机理是 DNA 的双股互补移码变异。

(4) 对于 PAH 致癌问题，应该考虑距离因素和传输因素，且亲电的双官能团的距离以 0.28 nm 为致癌的最优距离，而大于 6 个苯边单位的“双区”都没有活性。

(5) 采取忽略构型异构体间可能差别的近似，强调 PAH 代谢过程中酶的活性中心的空间要求的重要性。

双区理论提出了以 PAH 的离域能作参数的致癌活性定量公式(12.1)式：

$$\log K = 4.751\,\Delta E_1 \Delta E_2^{\frac{3}{2}} - 0.0512\, n \Delta E_2^{-3} \qquad (12.1)$$

式中，ΔE_1 和 ΔE_2 是具有最高和次高离域能的两活性中心的离域能，n 是脱毒区的数目。当

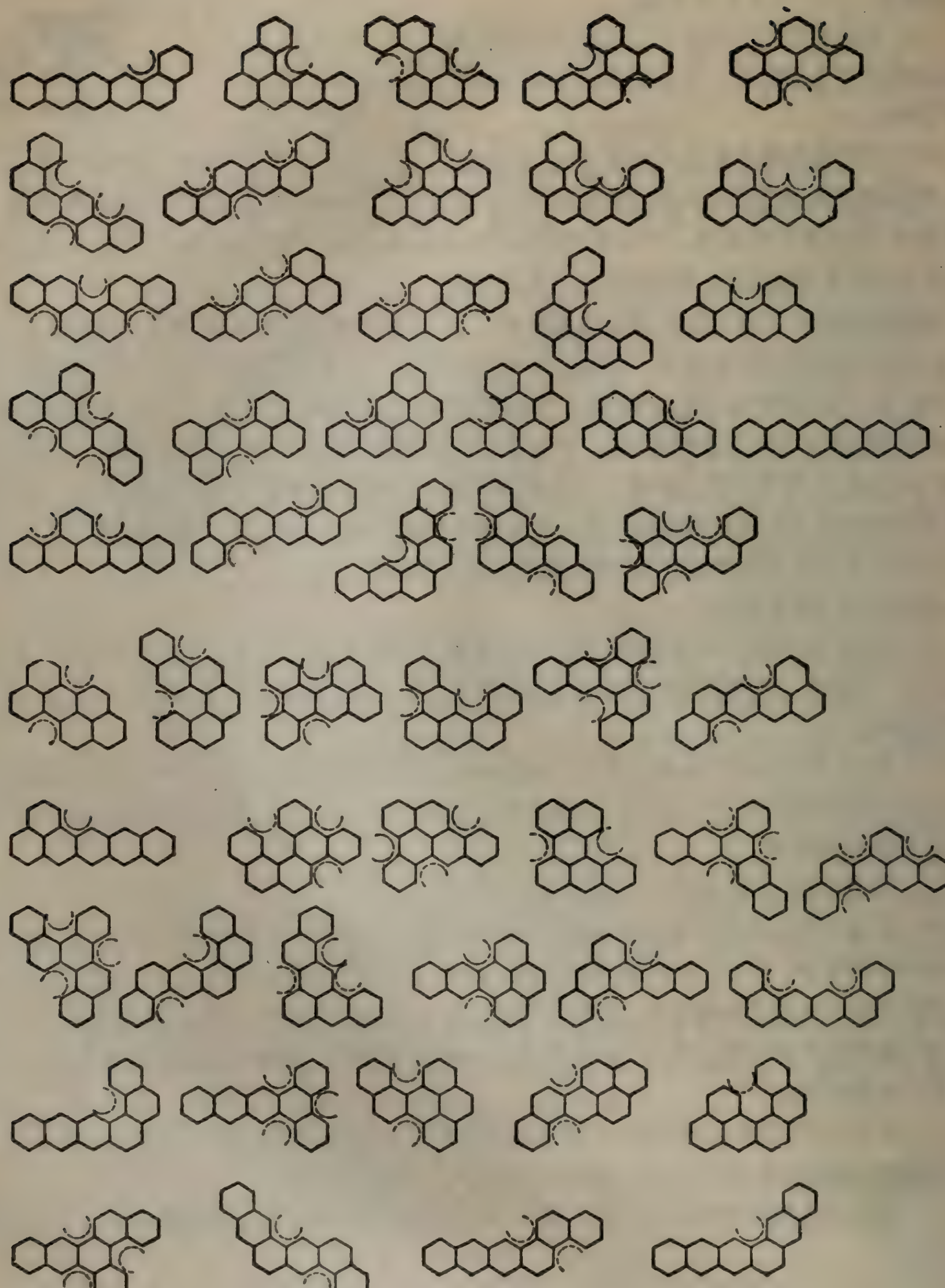
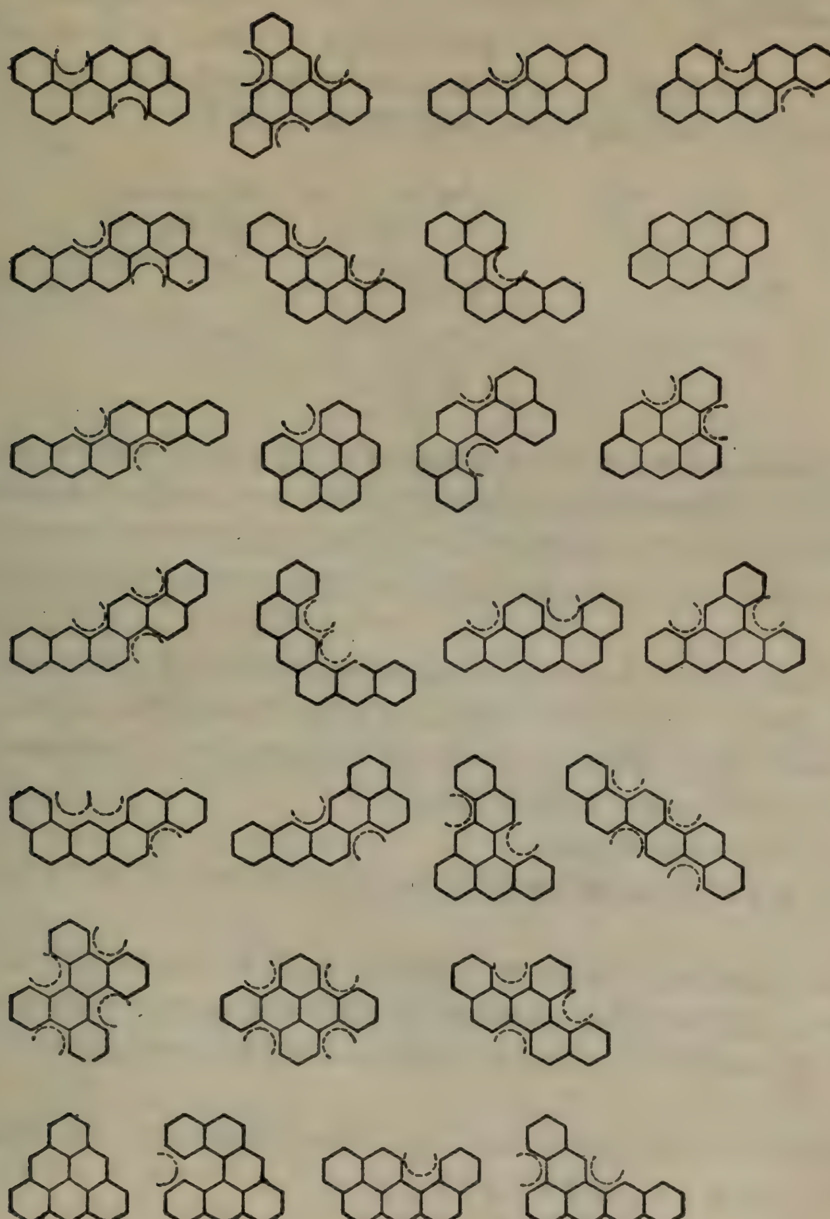


图 12-11 一些多



环芳烃的湾区[18]

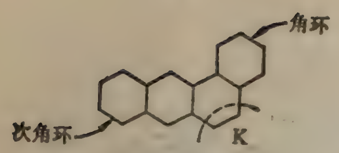
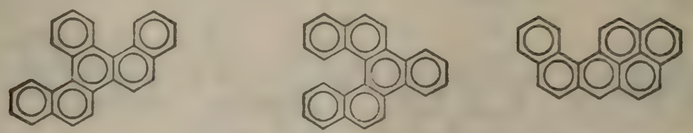


图 12-12 苯并[a]蒽的活性区域

$K<6$ 时,PAH 是不致癌的( );K 为6—15时,则是微弱致癌的(+); K 为 15—45 时,是致癌的(++);45—75 为显著致癌(+++),而 $K>75$ 者为强烈致癌(++++)。

戴乾圜运用(12.1)式计算了 49 个PAH,计算的致癌活性与实验结果的符合率达 98%(见表12-2)。不仅如此,他还预言了 4 个没有作过生物实验的 PAH 的致癌活性(图 12-13)。



$K\approx2.21(-)$　　$K\approx3.36(-)$　　$K\approx24.6(++)$

图 12-13

由于双区理论是建立在 PAH 体内代谢实验的基础上,且较合理地考虑了 PAH 分子中各关键区的作用。因此,计算的结果与实验有很高的符合率,对 PAH 的致癌活性也有预测能力。此外,戴乾圜等还对 36 个缺乏实验数据的六环 PAH 用双区定量公式进行了计算,对其致癌活性作了理论预测,指出其中的 9 个化合物可能是潜在的致癌剂。当然,这一结论是否正确,尚待实验证明。

综上所述,双区理论对 PAH 无论是计算的数量和结果都优于K区理论和湾区理论。双区理论提出的一些假设,可以认为,是基本合理的。但是,也有学者提出质疑,认为双区理论最根本的错误就是:量子模型与代谢模型不符,并且在双区公式的推导过程中,某些数据的处理不太严谨,目前也还没有足够的实验事实支持这个理论。

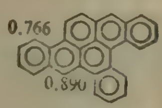
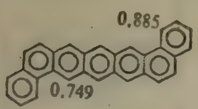
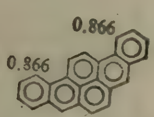
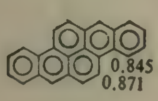
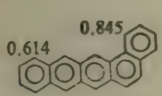
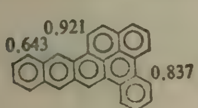
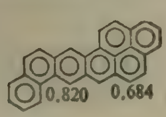
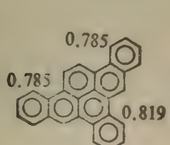
12.2.1.5 多环芳烃的构型效应

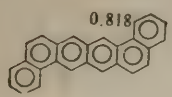
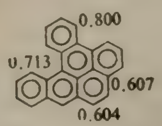
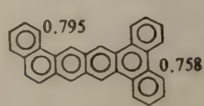
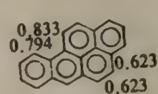
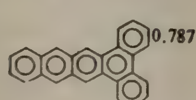
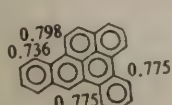
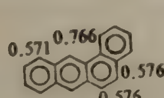
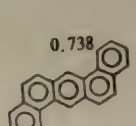
近几年来,在多环芳烃最终致癌物构型效应的量子生物学研究方面,取得了一些有意义的结果。

Jerina 等[20]实验证实,多环芳烃湾区二醇环氧化物为其最终致癌物(参见 12.1.3.2)。有顺反两种异构式,且异构式中由于羟基排列不同又有两种构型,即平伏构型和垂直构型(参见表 12-3)。进一步实验表明[21-25],无论顺式,还是反式,其平伏键型的致癌活性皆高于垂直键型的致癌活性。

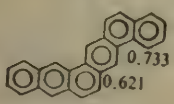
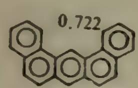
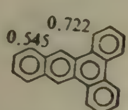
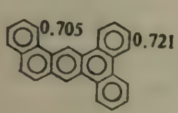
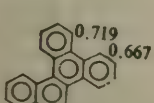
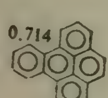
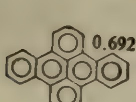
白迺彬等应用 CNDO/2 法研究了苯并[C]菲二醇环氧化物三个异构体(见表 12-3 中的顺式(eq,eq')-1,2-环氧-3,4-二醇苯并[C]菲 I,反式(eq,eq')-1,2-环氧-3,4-二醇苯并[C]菲II,顺式(ax,ax')-1,2-环氧-3,4-二醇苯并[C]菲III)以及三个模型化合物(顺式(eq,eq')-1,2-环氧-3,4-二醇环己烯 IV,反式(eq,eq')-1,2-环氧-3,4-二醇环己烯 V,顺式(ax,ax')-1,2-环氧3,4-二醇环己烯 VI)。计算表明,二醇环氧化物的三个异构体,以平伏键型异构式比垂直键型异构式具有更大活性。而计算三个模型化合物环己烯二醇环氧化物-氨系统的电子迁移特性(%)的结果表明,二醇环氧化物的构型是平伏键构型,还是垂直键构型,可能是决定生物活性的关键。

表 12-2

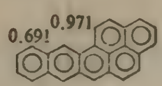
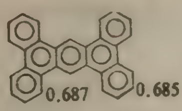
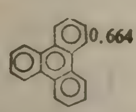
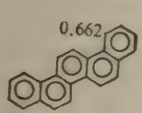
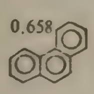
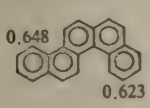
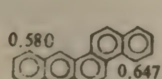
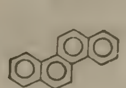
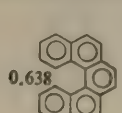
编号	化合物结构	化合物名称	湾区最高离域能 ΔE (β单位)	ΔE_1 (β单位)	ΔE_2 (β单位)	n	计算致癌活性指标 K	实验致癌活性	说明
1		三苯并[b,def,p]䓛	0.890	0.890	0.766	1	61.17 (+++)	(++)	
2		二苯并[a,l]戊省	0.885	0.885	0.400	3	0.01 (—)	(−)	取 $\Delta E_2=0.400\beta$
3		苯并[rst]戊芬	0.866	0.866	0.866	2	327 (++++)	(++++)/(+++)	
4		二苯并[b,def]䓛	0.845	0.845	0.871	1	376 (++++)	(++++)/(+++)	
5		苯并[a]丁省	0.845	0.845	0.614	2	3.07 (—)	(—)	
6		苯并[a]萘并[8,1,2-cde]丁省	0.837	0.921	0.643	2	6.00 (+)	(+)	
7		苯并[a]萘并[2,1,8-hij]丁省	0.820	0.820	0.684	1	12.21 (+)	(+)	
8		二苯并[h,rst]戊芬	0.819	0.785	0.785	2	39.2 (++)	(++)	

编号	化合物结构	化合物名称	湾区最高离域能 ΔE (β单位)	ΔE_1 (β单位)	ΔE_2 (β单位)	n	计算致癌活性指标 K	实验致癌活性	说明
9		三苯并[a,j]丁省	0.818	0.818	0.400	3	0.01 (−)	(−)	取$\Delta E_2=0.400\beta$
10		二苯并[def,p]䓛	0.800	0.800	0.713	1	17.20 (++)	(++)/ (+++)	
11		三苯并[a,c,j]丁省	0.795	0.769	0.758	3	17.32 (++)	(+)	
12		苯并[a]芘	0.794	0.794	0.833	1	123.55 (++++)	(++++)	
13		二苯并[a,c]丁省	0.787	0.787	0.787	3	32.17 (++)	(++)	
14		萘并[1,2,3,4-def]䓛	0.775	0.736	0.798	1	47.44 (+++)	(+++)/ (++++)	等价角环不被采用
15		苯并[a]蒽	0.766	0.766	0.571	2	1.34 (−)	(−)/ (+)	
16		二苯并[a,h]蒽	0.738	0.738	0.738	1	19.14 (++)	(++)	

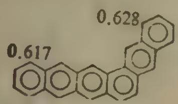
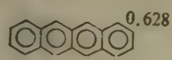
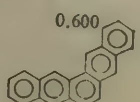
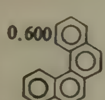
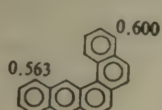
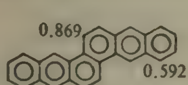
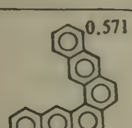
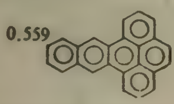
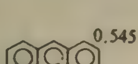
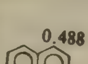
(续表)

编号	化合物结构	化合物名称	湾区最高离域能 ΔE (β单位)	ΔE_1 (β单位)	ΔE_2 (β单位)	n	计算致癌活性指标 K	实验致癌活性	说明
17		苯并[a]戊芬	0.733	0.733	0.621	2	2.55 (−)	(−)	
18		二苯并[a,j]蒽	0.722	0.722	0.722	1	14.29 (≈++)	(++)	
19		苯并[b]三亚苯	0.722	0.722	0.545	2	0.837 (−)	(−)/(+)	
20		萘并[1,2-b]三亚苯	0.721	0.721	0.705	2	8.09 (+)	(−)	
21		苯并[g]䓛	0.719	0.719	0.667	1	6.94 (+)	(+)	
22		苯并[e]芘	0.714	0.714	0.400	0	1.65 (−)	(−)	取$\Delta E_2=0.400\beta$
23		戊省	0.709	0.709	0.400	3	0.007 (−)	(−)	取$\Delta E_2=0.400\beta$
24		二苯并[fg,op]丁省	0.692	0.692	0.692	0	12.28 (+)	(+)/(−)	

(续表)

编号	化合物结构	化合物名称	湾区最高离域能 ΔE (β单位)	ΔE_1 (β单位)	ΔE_2 (β单位)	n	计算致癌活性指标 K	实验致癌活性	说明
25		萘并[8,1,2-cde]丁省	0.691	0.971	0.691	1	23.28 (++)	(++)	
26		菲并[9,10-b]三亚苯	0.687	0.687	0.685	3	3.73 (−)	(−)	
27		三亚苯 (Triphenylene)	0.661	0.664	0.664	1	5.61 (−)	(−)	
28		苉	0.662	0.662	0.662	1	5.45 (−)	(−)	
29		菲	0.658	0.658	0.658	1	5.14 (−)	(−)	
30		苯并[c]䓛	0.648	0.648	0.623	0	5.55 (≈+)	(++)	
31		苯并[b]䓛	0.647	0.647	0.580	1	2.11 (−)	(−)	
32		䓛	0.639	0.639	0.639	0	6.20 (+)	(+)	
33		二苯并[c,g]菲	0.638	0.638	0.638	1	3.89 (−)	(−)	

(续表)

编号	化合物结构	化合物名称	湾区最高离域能 ΔE (β单位)	ΔE_1 (β单位)	ΔE_2 (β单位)	n	计算致癌活性指标 K	实验致癌活性	说明
34		己芬	0.628	0.628	0.617	3	1.11 (−)	(−)	
35		丁省	0.628	0.628	0.628	2	2.12 (−)	(−)	
36		戊芬	0.600	0.600	0.600	3	0.80 (−)	(−)	
37		苯并[c]菲	0.600	0.600	0.600	0	6.20 (+)	(+)	
38		萘并[1,2-a]蒽	0.600	0.600	0.563	1	1.67 (−)	(−)	
39		二苯并[b,k]䓛	0.592	0.869	0.592	3	1.31 (−)	(−)	
40		蒽并[1,2-a]蒽	0.571	0.571	0.571	2	0.90 (−)	(−)	
41		二苯并[de,qr]丁省	0.559	0.559	0.400	1	0.23 (−)	(−)	取$\Delta E_2=0.400\ \beta$
42		蒽	0.545	0.545	0.545	1	1.27 (−)	(−)	
43		萘	0.488	0.488	0.488	0	1.86 (−)	(−)	

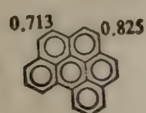
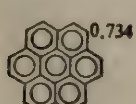
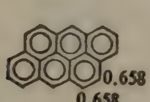
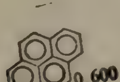
无角环区的化合物									
编号	化合物结构	化合物名称	湾区最高离域能 ΔE（β单位）	ΔE_1（β单位）	ΔE_2（β单位）	n	计算致癌活性指标 K	实验致癌活性	说　明
44		二苯并[cd,lm]苝		0.928	0.600	0	8.96 (+)	(+)	实际上取两个K区的不同离域能作参数
45		苯并[ghi]苝		0.826	0.713	0	26.45 (++)	(++)	
46		蔻(Coronene)					<1 (−)	(−)	除没有角环外，也无活性K区
47		二苯并[def,mno]䓛					<1 (−)	(−)	除没有角环外，也无活性K区
48		芘					<1 (−)	(−)	除没有角环外，也无活性K区
49		苝					<1 (−)	(−)	除没有角环外，也无活性K区

表 12-3 母体烃二醇环氧化物及模型化合物前线、近前线轨道能级表

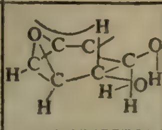
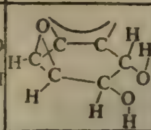
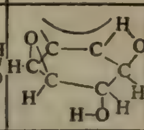
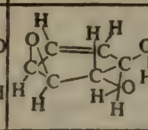
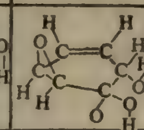
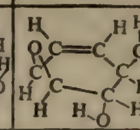
化合物	I	II	III	IV	V	VI
						
前线、近前线轨道能级	0.1814 0.1408 0.0755 0.0488	0.1824 0.1390 0.0734 0.0513	0.1930 0.1474 0.0854 0.0572	0.2565 0.2493 0.2350 0.1176	0.2579 0.2365 0.2328 0.1000	0.2681 0.2524 0.2351 0.1192

表 12-4 芳烃最终致癌物质构型效应预报与实验对照

类别	化合物	预报活性	实验			
			老鼠数	瘤/老鼠	有瘤鼠%	文献
I	反式(*ax*,*ax*′)-1,2-环氧-3,4-二醇苯并[e]芘	低	致癌活性低			[28]
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇苯并[e]芘	低				
	反式(*ax*,*ax*′)-1,2-环氧-3,4-二醇三亚苯	低				
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇三亚苯	低				
II	反式(*eq*,*eq*′)-1,2-环氧-3,4-二醇苯并[a]芘	高	29	1.50	43	[29]
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇苯并[a]芘	低	28	0.07	0.6	
	反式(*eq*,*eq*′)-1,2-环氧-3,4-二醇苯并[a]蒽	高		1.9	70	[22]
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇苯并[a]蒽	低		0.6	43	
	反式(*eq*,*eq*′)-1,2-环氧-3,4-二醇䓛	高	100	15.91	98	[24]
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇䓛	低	100	0.34	29	
	反式(*eq*,*eq*′)-1,2-环氧-3,4-二醇菲	高	100	0.16	16	
	顺式(*ax*,*ax*′)-1,2-环氧-3,4-二醇菲	低	100	0.18	14	
III	反式(*eq*,*eq*′)-1,2-环氧-3,4-二醇苯并[c]菲	高	30	8.12	80	[26] [27]
	顺式(*eq*,*eq*′)-1,2-环氧-3,4-二醇苯并[c]菲	高	30	7.03	89	

通过计算，所得的结果(表 12-3 和表 12-4)与文献[21—29]所报导的实验结果相符。由此可见,PAH 最终致癌物二醇环氧化物的构型确对分子特性有影响，从而对致癌活性产生影响。其中顺反异构影响不大,但环氧对位羟基是取平伏,还是取垂直键型影响较大。平伏构型致癌活

性大于垂直构型[32]。

就计算构型效应所用的方法来看，过去应用 PMO 法计算PAH二醇环氧化物形成三醇碳阳离子之离域能 $\Delta E/\beta$，系按平面模型进行，不能反映出构型效应，因而往往计算活性指标顺序与实验顺序有所不符(参见表 12-5)。

表 12-5 PMO 与 CNDO 计算预报对照表

芳 烃	苯并[a]芘	苯并[e]芘	三亚苯	䓛	苯并[c]菲	文 献
Iball 指标*	73	2	0	3	4	[30]
PMO($\Delta E/\beta$)	0.794	0.714	0.664	0.639	0.600	[31]
构型(主体)	平伏、垂直	垂直	垂直	平伏、垂直	平伏	[22,24,26—29]

* 癌发生率的指标。

继白逎彬等[32]的工作之后，陈凯先等[33]应用从头计算(STO-3G)和分子静电势方法研究了PAH 最终致癌物致癌活性与其空间构型的关系。从计算所得的 E_{LUMO}值、环氧与碳原子间重迭集居数、偶极矩以及分子静电势图的特征分析(表 12-6 和图 12-14—18 所示的模型分子三种异构体)，表明无论是顺式或反式构型，其二醇羟基取平伏键均比取直立键构型的致癌活性强。这与白逎彬等的理论计算和上述文献所报导的实验事实相符，从而在 PAH 最终致癌物的构型效应上支持了 Jerina 的湾区理论。

表 12-6 模型分子三种异构体的电子性质

	I 反式(eq,eq')	II 顺式(ax,ax')	III 顺式(eq,eq')
总能量(a.u.)	−450.5140	−450.5064	−450.5072
轨道能量(a.u.)			
34(HOMO)	−0.2772	−0.2688	−0.2848
36(LUMO)	0.2599	0.2637	0.2601
重迭集居数			
$P_{C_3-O_7}$	0.3981	0.3974	0.3951
$P_{C_4-O_7}$	0.4061	0.4032	0.4070
$\Delta P=P_{C_4-O_7}-P_{C_3-O_7}$	0.0080	0.0058	0.0119
净电荷			
C_3	0.0195	0.0213	0.0197
C_4	0.0218	0.0235	0.0171
O_7	−0.2265	−0.2137	−0.2311
偶极矩 (debye)	2.2775	1.2926	2.0057
分量* x 轴	1.5317	−0.0991	0.8839
y 轴	−1.5237	−0.3073	−0.9596
z 轴	−1.0435	−1.2516	−1.5234

* x 轴通过C_1—C_2的联线；C_1、C_2、C_3、C_5 原子在 xy 平面上。

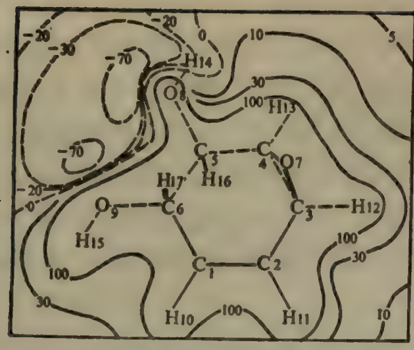


图 12-14 反式(eq,eq')异构体:分子静电势图

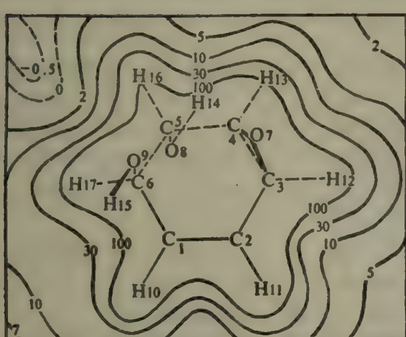


图 12-15 顺式(ax,ax')异构体:分子静电势图

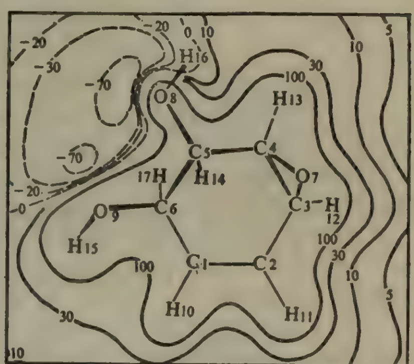
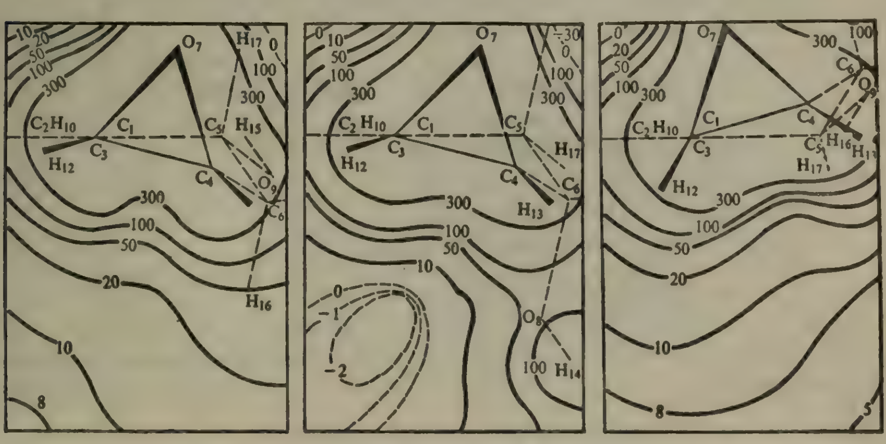


图 12-16 顺式(eq,eq')异构体:分子静电势图



I 反式(eq,eq')　　II 顺式(ax,ax')　　III 顺式(eq,eq')

图 12-17 垂直平面上的分子静电势

12.2.1.6 多环芳烃致癌能力的参数图

研究的实践提示我们,为了关联 PAH 的致癌能力,不但要考虑关键区域的电子因素,还要同时考虑几何因素。在 Berger 和 Herndon 工作的基础上,陈念贻等[34]尝试用参数图分析的方

法,对 58 个PAH进行致癌能力的区域划分,进一步探讨了湾区理论。

对已做过致癌能力试验的 58 个 PAH,其中有 22 个化合物注射到试验动物上产生肉瘤,或者对试验动物进行皮试时,产生乳腺癌和皮癌。特别其中有 10 个致癌化合物研究较为仔细,且它们的致癌能力较强,其 Iball 指标列于表 12-7。当我们分析这 58 种PAH的分子结构时,首先看到有 14 种化合物没有湾区,它们之中绝大部分没有致癌能力,少部分即使有致癌能力,也不明显。另外有三种化合物,四苯并萘、二苯并[c,g]菲、二苯并[b,g]菲,虽有湾区,但因存在较大空间障碍,使湾区致癌历程难于进行,也没有致癌能力。余下 41 种化合物的致癌能力的参数示于图 12-18。

表 12-7 多核芳烃致癌活性区域分布表

无 Bay 区		(－) 区				(＋) (++) 区				(+++) (++++) 区**	
化合物	活性	化合物	*P*	*S*	活性	化合物	*P*	*S*	活性	化合物	活性
萘	(－)	三亚苯	0.656	108	(－)	菲	0.600	79	(＋)	苯并[a]芘	(++++)
蒽	(－)	二苯并[e,l]芘	0.526	126	(－)	䓛	0.625	95	(＋)	二苯并[a,l]芘	(++++)
丁省	(－)	苯并[a]丁省	0.556	127	(－)	苯并[c]菲	0.625	98	(＋)	二苯并[a,h]芘	(++++)
戊省	(－)	二苯并[a,c]蒽	0.546	134	(－)	苯并[a]蒽	0.571	106	(＋)	二苯并[a,i]芘	(++++)
蒽[2,1-a]蒽	(－)	二苯并[a,l]丁省	0.563	137	(－)	二苯并[a,j]蒽	0.583	116	(＋)	二苯并[a,e]芘	(+++)
芘	(－)	三苯并[a,c,j]蒽	0.563	142	(－)	苯并[c]䓛	0.615	116	(＋)		
萘[2,3-e]芘	(－)	四苯并[a,c,h,i]蒽	0.550	145	(－)	苯并[e]芘	0.545	118	(＋)	苯并[b]䓛	(－)
萘[2,3-a]丁省	(－)	苯并晕苯	0.588	145	(－)	苯并[g]䓛	0.571	119	(++)		
戊菲	(－)	苯并[a]戊省	0.545	148	(－)	二苯并[a,h]蒽	0.583	119	(++)		
芘	(＋)	二苯并[a,j]丁省	0.563	149	(－)	菲并[2,3-b]芘	0.572	126	(＋)		
晕苯	(＋)	二苯并[a,h]芘	0.750	157	(－)	萘并[2,3-b]芘	0.750	129	(++)		
苯并[1,12]芘	(＋)	二苯并[a,h]戊省	0.550	158	(－)	三苯并[a,e,j]芘	0.630	165	(++)		
蒽垛蒽	(＋)	二苯并[a,c]丁省	0.529	158	(－)						
佩芘	(＋)	苯并[c]戊菲	0.588	158	(－)	苉	0.615	110	(－)		
		二苯并[a,l]戊省	0.550	171	(－)	皮蒽	0.652	150	(－)		
四苯并萘*	(－)	四苯并丁省	0.571	171	(－)						
二苯并[c,g]菲*	(－)	三苯并[a,e,j]丁省	0.583	171	(－)						
二苯并[b,g]菲*	(－)	二苯并[b,k]芘	0.714	172	(－)						
		二苯并[1,2-b,k]䓛	0.542	179	(－)						
		四苯并[a,e,hi,gr]戊省	0.636	184	(－)						
		二苯并[2,3-bh]芘	0.731	188	(－)						
17	12	21			21	21			19	6	5

* 分子 Bay 区有空间障碍。

** P、S 数值见表 3。

图 12-18 纵坐标 S 是 PAH 分子参加反应的障碍面积 (incumbrance area)。在计算时,应用了以下参数:PAH相邻碳原子距离为 1.41 Å;芳环间在单键相连时,C—C 键长取 1.54 Å;芳环上C—H基的 Van der Waals 直径为 3.70 Å。为了易于比较,计算时所有 PAH 皆取平面。图中,横坐标 P_B 是湾区的 Paulingπ 键级,为表征湾区键定域能的键参数。

由图 12-18 可以明显看出，PAH 的致癌能力依其结构不同，分为三个区域。致癌能力 Iball 指标大于 50 者居于中央；致癌能力 Iball 指标在 0—50 之间的，在它的外围，无致癌能力的在更外围。大致是：由湾区 Pauling π 键级 $P_0=0.667$，$S_0=120$ Å² 向外，沿辐射方向，致癌能力由强变弱，到失去致癌能力，呈现十分有趣的分布。我们将实验活性数据和图 12-19 分布点对应列于表 12-7 中。对于 58 个 PAH，有出入的有 8 个，占 1/7，出入较大的，仅有苯并[b]䓛、皮蒽、苉三个化合物，皆较前人的划分为好。

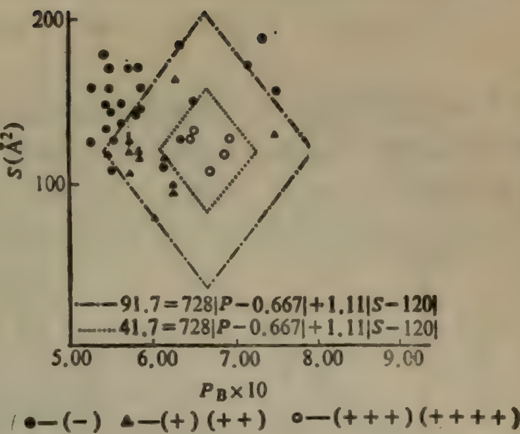


图 12-18 多核芳烃致癌能力键参数图

●——(−)，▲——(+)(++)，

■——(+++)(++++)

12.2.1.7 湾区 π 键级、分子大小与致癌能力关系

基于图 12-18 的致癌能力分布情况，我们有理由假设，在 $P_0=0.667$，$S_0=120$ Å² 区附近存在一假想的化合物 A，它的致癌能力最大，若其致癌反应速度常数为 K_0，则：

$$K_0=\frac{KT}{h}\exp[-G_0^{\ast}/RT]$$

$$=\frac{KT}{h}\exp[-\Delta H_0^{\ast}/RT]\exp[\Delta S_0^{\ast}/R] \tag{12.2}$$

若有一 PAH B 与 A 同时进入人体，将与体内的 NADPH 和 O_2 产生相似平行竞争反应，生成各自的最终致癌物质。用下式表示：

$$A+NADPH+O_2\xrightarrow{K_0}\text{最终致癌物质A}'$$

$$B+NADPH+O_2\xrightarrow{K}\text{最终致癌物质B}'$$

令 K 为第二个反应常数：

$$K=\frac{KT}{h}\exp[-G^{\ast}/RT]$$

$$=\frac{KT}{h}\exp[-\Delta H^{\ast}/RT]\exp[\Delta S^{\ast}/R] \tag{12.3}$$

求(12.3)/(12.2)，则有：

$$K/K_0=\exp[-\Delta G^{\ast}/RT+\Delta G_0^{\ast}/RT] \tag{12.4}$$

进一步对方程(12.4)进行级数展开，作为一级近似，取其前两项，得：

$$K/K_0=1-(\Delta G^{\ast}-\Delta G_0^{\ast})/RT$$

$$=1-[\Delta H^{\ast}-\Delta H_0^{\ast}]/RT+[\Delta S^{\ast}-\Delta S_0^{\ast}]/R \tag{12.5}$$

式中活化焓一项一般与能量有关。对于我们所考虑的反应体系，可取其正比于湾区的定域能或 π 键级。活化熵项对于我们所考虑的体系与几何因素有关，可取其正比于分子障碍面积。对于一

系列相似反应,比例数可视为近似相等,于是(12.5)式经整理为:

$$K=K_0-a\Delta P_B-b\Delta S$$
$$=K_0-a|P-P_0|-b|S-S_0| \quad (12.6)$$

式中 a、b 皆为正的常数。当取 Iball 致癌能力指标正比于上述反应速度常数 K,并将 10 个经过较仔细实验研究的 PAH 之 Iball 值,P_B、S 值依式(12.6) 做最小二乘法线性拟合, 得到 a、b 和 K_0值,即有:

$$\text{Iball}=91.7-728\Delta P_B-1.11\Delta S \quad (12.7)$$

当 Iball 为 50 和 0 时,即为图 12-18 中分界线方程。

式(12.7)的计算值与实验值并列于表 12-8, 示于图 12-19。由表 12-8 可看到符合良好。经计算,实验值与计算值之相关系数达到 0.975。应该指出,皮蒽、芘和苯并 [b] 蔗, 根据湾区理论分子在湾区,能量上也接近 $P_0=0.667$, 分子大小接近 120 Å^2, 似乎应有较强的致癌活性。现将根据(12.7)式计算的 Iball 值亦列入表 12-8, 以期引起实验工作者的注意。

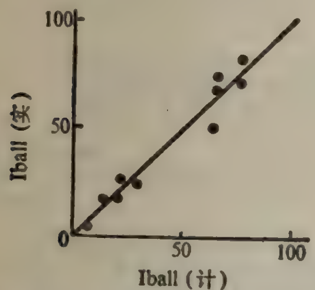


图 12-19 多核芳烃致癌能力计算值实验值相关性图

表 12-8 多核芳烃致癌能力计算值与实验值比较表

化合物	P_B	S Å^2	Iball (实)	Iball(计)
苯并[a]蒽	0.571	106	5	6
三苯并[a,e,j]芘	0.630	165	17	15
苯并[g]蔗	0.571	119	18	21
二苯并[a,h]蒽	0.583	119	26	29
苯并[2,3-b]芘	0.750	129	27	22
二苯并[a,e]芘	0.647	132	50	64
二苯并[a,h]芘	0.692	127	68	66
苯并[a]芘	0.667	106	72	76
二苯并[a,i]芘	0.643	129	74	65
二苯并[a,l]芘	0.687	119	82	77
苯并[b]蔗	0.636	127	(—)	60
芘	0.615	110	(—)	43
皮蒽	0.651	150	(—)	43

由 (12.7) 式来看,a 大于 b, 说明反应的电子因素所起的作用比几何因素所起的作用为重要,只有当分子过大或过小时,第三项才是有意义的因素。

从以上讨论可以看到,致癌反应不在定域能更小的K区,而在湾区进行, 本身是一个耐人寻味的理论问题。随着 PAH 致癌作用的深入研究,这个谜将会逐步揭开。

12.2.1.8 多环芳烃的定量结构致癌活性关系

关于结构致癌活性关系的理论研究,主要集中于对 PAH 的研究, 而引起人们关注的一些理论,总是采用量子化学的方法来讨论这一问题。

PAH 的代谢过程是致癌活化与降解脱毒两方面的综合代谢过程。为了使读者对 PAH 的定量结构致癌活性关系的理论研究有一个初步的了解，以下我们将主要介绍戴乾圜等[6,12]的工作。

按照戴乾圜等的双区理论，如图 12-20 所示，角环和次角环是 PAH 分子中突出的环，前者与分子轴近乎垂直，后者与分子轴近乎平行。由于角环及次角环的有利几何形态，易于进入单氧酶活性中心所处的缝隙。PAH分子在形成M区二氢二醇后，PAH 的溶解度将增加，因此，这对其向细胞核的转移是有利的。根据分子轨道理论和组合论数学，可证角环 E 碳的离域能总是高于

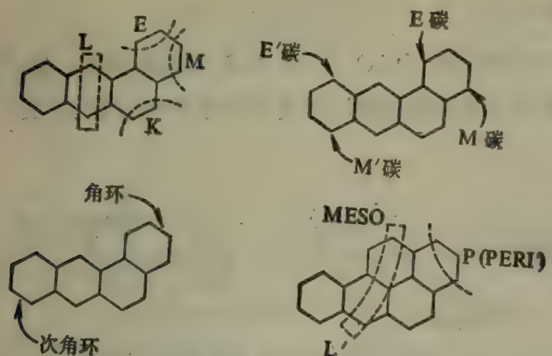


图 12-20

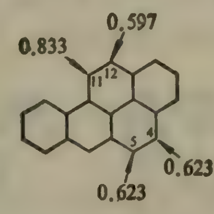


图 12-21

其它角环或次角环碳原子的离域能。类似地，可证 L 区的碳原子总是具有最低的定域能，而在K区上，往往但并非总是具有次低的定域能。由于上述原因，角环的代谢趋势及其与 DNA 结合的趋势(参见 12.3.1)，往往超过由其定域能所估计的，于是，角环代谢成一强力的烷化剂。因此，角环 E 碳是一重要的活性区域，这与实验相符合。从几何形态及溶解度方面而论，次角环与角环是类似的，但是 E′碳的离域能远低于 E 碳。因此，次角环也是一个活性区域，但只具有次要的重要性。戴乾圜等在系统地计算了各种 PAH K 区碳阳离子的离域能以后，发现有两类极端的情况。例如苯并[a]芘的 11 位碳上的碳原子具有非常高的离域能 (0.833 β)，但在 12 位碳上则具有远为较低的离域能(0.597 β)，如图 12-21 所示。因此，苯并[a]芘的 11，12-环氧化物易于分解形成 11 位碳上的碳阳离子，并转化成一个烷化剂。这类 K 区称为活性 K 区，这是一个活性区域和亲电中心。与之相反，4,5 位上的 K 区则两碳原子的离域能相等(均为 0.623 β)。这类K区称为等价K区，它易于转化产生二氢二醇，但却较难形成碳阳离子，因此，这类 K 区主要是一个脱毒区。L 区的两个碳原子与其它位置比较，总是具有最低的定域能，因此，L 区易于氧化形成醌。醌的致癌活性皆远低于母体化合物。L区是一个脱毒区，P 区是周稠型 PAH 所专有的区域。P 区的环氧化与其它区域的环氧化不同，其它区域的环氧化只引起一个苯环的定域，而 P 区的环氧化则引起两个苯环的定域。P 区环氧化物的形成，在能量上是不利的，它立即借 NIH 移位重排成酚，这是一个脱毒过程。但另一方面，其中具最低定域能的位置，可与 DNA 分子结合，这可能是经 P 区环氧化物实现的，这是一种活化过程(图 12-22)。

由于 P 区贡献的两重性，故可忽略所有 P 区的影响。根据以上分析，戴乾圜等发现，对于

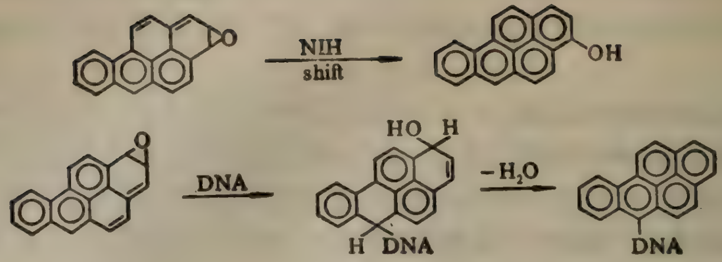


图 12-22

PAH 的致癌活性的必要条件，是存在着两个活性区域（中心），图12-23所列可为例证。两个活性中心对于 PAH 的活性是足够的，若有任何第三个多余的角环，则只会加速脱毒过程，因此，多余的角环是一脱毒区，如图12-24。

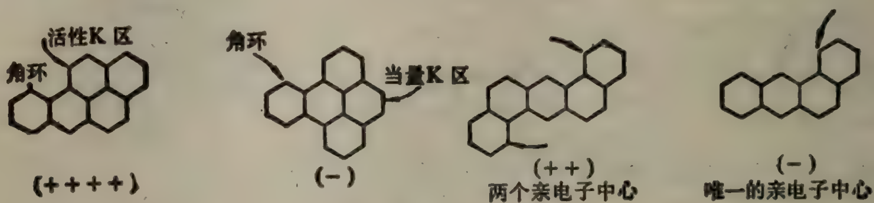


图 12-23

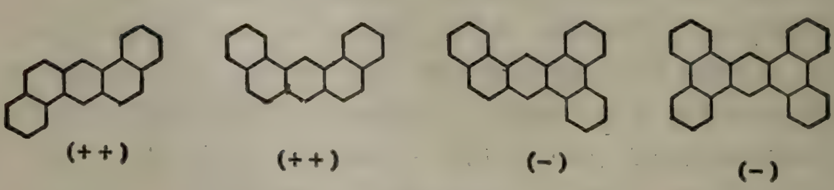


图 12-24

成对亲电碳原子的最优距离大约为 2.80 —3.00Å，从图12-25所列的结构式中可以看出，两个计算的活性中心，两者彼此都相隔 2.78 Å。在致癌的 PAH 中，也有比 2.80—3.00 Å 的距离大的例证，如图12-26所示。若 PAH 分子中两活性中心的距离超过 6 个苯边单位时，其致癌活性将丧失，如图12-27所列。

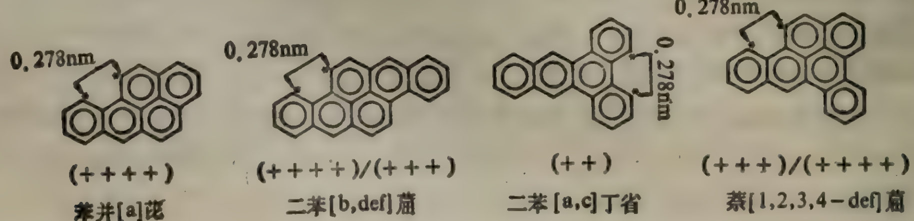
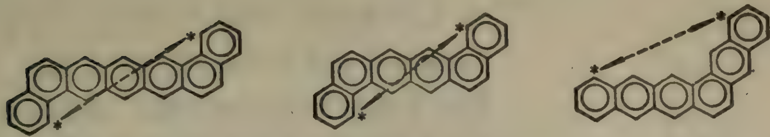


图 12-25



(++)　(++)/(+++)　(++)

图 12-26



(−)　(−)　(−)

图 12-27

戴乾圜等应用式(12.1)(参见12.2.1.4)所作的理论计算表明,计算的致癌活性与实验比较一致,实验与理论的符合达98%

迄今,有关结构生理活性关系的研究,仍然是一个疑难的科学问题。但是,结构致癌活性关系的研究已经表现出在阐明致癌作用的机理方面,将是一个强有力的工具,因而在这一领域吸引了愈来愈多的科学工作者。

12.2.2 甲基多环芳烃

大量的研究说明,甲基基团在 PAH 的引入, 将通过两个一般方式——空间效应和电子效应,对PAH 的致癌活性产生复杂的影响。与此同时,甲基本身也会因为这样两个效应而受到影响。

12.2.2.1 甲基的空间效应和电子效应

1. 空间效应

甲基对 PAH 的影响,首先是产生空间效应。这种效应又与甲基的大小和位置有关。由于甲基具有较大的空间体积,因而对分子中拥挤程度的影响是显而易见的。Pople 等用 INDO 法计算了甲基的平衡几何构型,结果与实验数据基本一致(表12-9)。且 INDO 和CNDO 计算的平衡键角是相同的[35]。Lathan 等人[36]用 STO-3 G 从头计算法计算甲基的键长和键角也得到了类似的结果。

表 12-9

键长(Å)			键角(度)		
计算		实验	计算		实验
INDO	STO-3G		INDO	STO-3G	
1.110	1.080	1.079	119.0	118.3	120

就甲基的空间效应对 PAH 结构致癌活性关系影响的研究而言，只取得了一些经验的规律，迄今仍未找到普遍适用于各种 PAH 的统一规则。

在一些情况下，甲基引入到 PAH 的角环附近，对角环的酶催化代谢会产生阻碍作用，邻近脱毒区则阻滞脱毒。如果甲基的取代发生在角环上，不仅阻滞角环的活化，同时也影响到角环碳阳离子与生物大分子的共价结合，从而降低 PAH 的致癌活性。

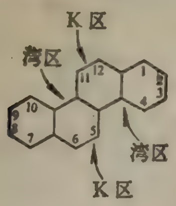


图 12-28

在另一些情况下，如䓛及其甲基衍生物的实验研究表明，5-甲基䓛是一种强烈的致癌剂，而䓛及其它单甲基衍生物则仅仅具有中等的或微弱的致癌性。从图 12-28 可以看出，䓛有两个 K区和两个湾区(参见表12-1)，其中 5 位和 11位甲基取代衍生物应该具有相同的致癌活性（当然亦包括其它对应位置的甲基取代衍生物）。我们的计算提示了这一点(表12-10)。但 Berger 等计算的理论反应活性指数

表 12-10 PAH 和甲基取代 PAH 的不同位置 S_r' 与致癌性的比较

化合物	S_r'			致癌性	
	K 区	E 区	M 区	皮肤	皮下
苯并[a]蒽	0.760	0.181	0.153	−	−
1-甲基苯并[a]蒽	0.758	0.181	0.152	−	−
2-甲基苯并[a]蒽	0.806	0.210	0.177	−	−
3-甲基苯并[a]蒽	0.734	0.194	0.138	−	−
4-甲基苯并[a]蒽	0.842	0.197	0.162		
5-甲基苯并[a]蒽	0.928	0.188	0.157	−	+
6-甲基苯并[a]蒽	0.923	0.223	0.187	+++	+++
7-甲基苯并[a]蒽	0.787	0.186	0.160	++++	+++
苯并[a]芘	0.815	0.464	0.429	++++	++++
2-甲基苯并[a]芘	0.809	0.465	0.430		+++
3-甲基苯并[a]芘	0.863	0.462	0.428		+++
4-甲基苯并[a]芘	0.952	0.473	0.437		+++
7-甲基苯并[a]芘	0.800	0.477	0.496		++
9-甲基苯并[a]芘	0.809	0.533	0.414		−
12-甲基苯并[a]	0.780	0.473	0.438		+++
苯并[ah]蒽	0.784	0.212	0.169	+++	+++
14-甲基苯并[ah]蒽	0.798	0.233	0.186		
1-甲基苯并[ah]蒽	0.826	0.258	0.189	+	
4-甲基苯并[ah]蒽	0.946	0.222	0.182		
5-甲基苯并[ah]蒽	0.947	0.255	0.207	++	
6-甲基苯并[ah]蒽	0.815	0.210	0.172		+++
䓛	0.782	0.433	0.391	+	
5-甲基䓛	0.925	0.465	0.413	+++	++
6-甲基䓛	0.905	0.429	0.397		+
9-甲基䓛	0.808	0.520	0.374	−	
10-甲基䓛	0.780	0.533	0.482		+

却表现出不同的情形。

按照分子模型，5-甲基䓛的甲基基团跟6位氢原子间应有空间位阻，且这种位阻效应迫使整个分子偏离了平面，引起平面构型的变形。Hecht等指出，在甲基䓛系列中，从结构上导致致癌作用的充分条件是：一个湾区甲基和一个邻近未取代角环的自由迫位。按照Hecht等的观点，5-甲基䓛即具备这样的充分条件。他们认为甲基在湾区的存在是十分必要的，因为所有的单甲基䓛异构体都有一个自由角环和邻近的迫位，但具有强致癌活性的只有5-甲基䓛。这种在结构上导致致癌作用的充分条件也存在于其它一些分子中，如图12-29（有关甲基的空间效应可参见12.2.6.2）。

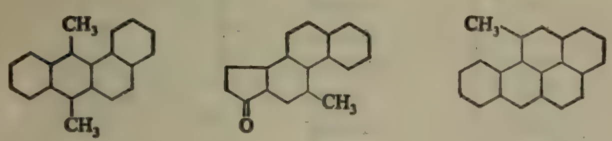


图 12-29

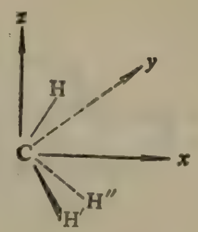


图 12-30

2．电子效应

迄今为止，有关取代基的电子效应，除了肯定至少有两种因素（诱导效应和共振效应）在起作用外，还没有得到一致公认的结论。

就甲基（$-CH_3$）而言，其坐标系如图12-30所示，其诱导效应跟它是“供电子基”还是“吸电子基”密切相关。实验资料说明，甲基所影响的许多种反应的方式与它的“供电子性”是相一致的。但是，也有关于甲基“吸电子性”的报导。近些年来，围绕“甲基是供电子基还是吸电子基”的问题展开了讨论。在我国，袁云程[37]、曾广植[38]、曾昭氘等[39]就此问题发表了讨论文章。陈念陔等用自洽场分子轨道从头计算法研究烷烃本身的电荷分布，针对甲基是“供电子基”还是“吸电子基”的问题，提出了他们的解释[40]。

我们计算PAH分子中各原子的净电荷、亲核超离域度 S_r^N（表12-11）的初步印象是：

（1）当甲基取代了PAH分子中的氢原子后，与甲基直接相连的碳原子上的电荷密度降低；与该碳原子邻接的两个碳原子的电荷密度增高。

（2）甲基取代的PAH跟母体PAH比较，前者分子中各原子（计算未涉及氢原子）的亲核超离域度 S_r^N 均有不同程度的降低。

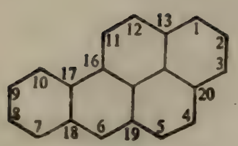
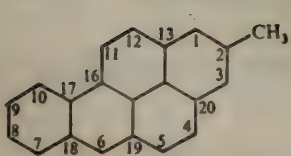
就以上HMO法的计算结果分析，甲基对于PAH分子是供电子基。

除诱导效应外，甲基的“超共轭效应”的真伪如何尚属有争议的问题，还有待于实验事实的进一步验证。不过，迄今有关“超共轭效应”的概念在解释某些现象时，仍具有它一定的价值。

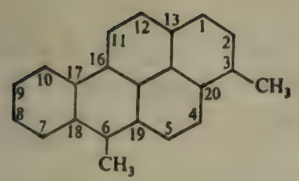
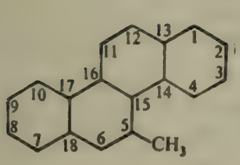
12.2.2.2 甲基取代多环芳烃的致癌活性

如前所述，湾区二醇环氧化物是一些PAH的终致癌形式。由图12-7和图12-21可以看到，M碳对苯角环的致癌代谢似乎起着关键作用。代谢一经发生，苯角环的共轭性被破坏，M区首先发生环氧化，并在环氧化物水解酶的作用下转化为二氢二醇化合物，接着是E区发生环氧化，

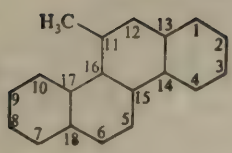
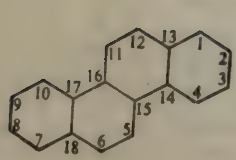
表 12-11　PAH分子中各原子的S_r^N和净电荷(HMO法)

分　　　子	原　子	S_r^N（负值）	净　电　荷
	1	1.1757	
	2	0.8340	
	3	1.1535	
	4	1.0699	
	5	1.0700	
	6	1.4079	
	7	1.0758	
	8	0.8687	
	9	0.9636	
	10	0.9704	
	11	0.9674	
	12	1.0948	
	13	0.7335	
	14	—	
	15	—	
	16	0.8795	
	17	0.8331	
	18	0.7002	
	19	0.7793	
	20	0.7797	
	1	1.1559	−0.0193
	2	0.8150	0.0397
	3	1.1449	−0.0167
	4	1.0693	−0.0010
	5	1.0699	0.0001
	6	1.4077	0.0002
	7	1.0758	0.0001
	8	0.8675	−0.0003
	9	0.9636	0.0000
	10	0.9693	−0.0003
	11	0.9674	0.0000
	12	1.0907	−0.0019
	13	0.7327	0.0009
	14	—	—
	15	—	—
	16	0.8751	−0.0016
	17	0.8331	0.0000
	18	0.6992	−0.0005
	19	0.7785	−0.0007
	20	0.7789	0.0009

(续表)

分　　子	原 子	S_r^N (负值)	净 电 荷
	1	1.1291	−0.0038
	2	0.8254	−0.0164
	3	1.0853	0.0393
	4	1.0259	−0.0067
	5	1.0376	−0.0022
	6	1.3308	0.0477
	7	1.0603	0.0030
	8	0.8590	−0.0033
	9	0.9508	0.0002
	10	0.9594	−0.0017
	11	0.9485	−0.0041
	12	1.0515	−0.0078
	13	0.7178	−0.0076
	14	—	—
	15	—	—
	16	0.8395	−0.0112
	17	0.8210	−0.0005
	18	0.6964	−0.0100
	19	0.7447	−0.0171
	20	0.7563	−0.0130
	1	0.9922	0.0000
	2	0.8631	−0.0007
	3	0.8969	0.0003
	4	0.9535	−0.0012
	5	0.9308	0.0436
	6	1.0025	−0.0282
	7	0.9799	−0.0048
	8	0.8642	0.0000
	9	0.8846	−0.0045
	10	0.9537	0.0005
	11	0.9519	−0.0003
	12	1.0430	0.0007
	13	0.7192	−0.0007
	14	0.7718	0.0015
	15	0.8265	0.0009
	16	0.8274	−0.0086
	17	0.7621	−0.0060
	18	0.7193	0.0012

（续表）

分子	原子	S_r^N（负值）	净电荷
	1	0.9799	−0.0048
	2	0.8642	0.0000
	3	0.8846	−0.0045
	4	0.9537	0.0005
	5	0.9519	−0.0003
	6	1.0430	0.0007
	7	0.9922	0.0000
	8	0.8631	−0.0007
	9	0.8969	0.0003
	10	0.9535	−0.0012
	11	0.9308	0.0436
	12	1.0025	−0.0282
	13	0.7193	0.0012
	14	0.7621	−0.0060
	15	0.8274	−0.0086
	16	0.8265	−0.0009
	17	0.7718	0.0015
	18	0.7192	−0.0007
	1	0.9923	
	2	0.8642	
	3	0.8972	
	4	0.9542	
	5	0.9535	
	6	1.0437	
	7	0.9923	
	8	0.8642	
	9	0.8972	
	10	0.9542	
	11	0.9535	
	12	1.0437	
	13	0.7204	
	14	0.7732	
	15	0.8274	
	16	0.8274	
	17	0.7732	
	18	0.7204	

使苯角环变成不为水化酶水解的二醇环氧化物,然后才在 C 10 位上形成亲电的碳阳离子(M区、E 区的位置见图12-31)。有文献称,[41]E 碳阳离子的浓度,可因甲基对 M 碳的影响而增加。

我们用 HMO 法对 54 个 PAH 和甲基取代 PAH 的K区,苯角环的M区和E区近似超离域度(S'_r)的计算结果显示: K区具有最大的 S'_r,其次是E区,M区最小。但是,当甲基引进 PAH 后,如果M区的 S'_r大于母体化合物M区的 S'_r,则致癌作用一般有所增强,且由于甲基对M碳的影响,致使E区 S'_r 增大。如果甲基取代在M区,则E区 S'_r 的增高更明显(参见表 12-10)。可是,从表 12-10 并不能得到有关近似超离域度与致癌活性的哪怕是较为明显的对应关系。然而,在探讨 PAH 及其甲基取代衍生物的结构致癌活性关系时,我们以$\sum S_r^E-\sum S_r^N$ 值为指标,获得了较满意的结果,其中尤其是甲基苯并[a]蒽取代衍生物的$\sum S_r^E-S_r^N$值与文献[11]的实验致癌活性有着很好的对应关系(表 12-12)。表 12-13 则列出采用 $\sum S_r^E-S_r^N$ 值获得的结果与 Memory[42]

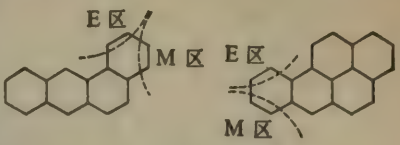


图 12-31 PAH 角环的 M 区和 E 区

表 12-12 单甲基苯并[a]蒽取代衍生物的$\sum S_r^E-\sum S_r^N$值与致癌活性的关系

苯并[a]蒽取代衍生物	$\sum S_r^E-\sum S_r^N$	作用部位	动物数	诱发肿瘤数
7-甲基	0.6436	皮肤 皮下	20 20	17(2个肉瘤) 9
12-甲基	0.5830	皮肤 皮下	20 20 60(大鼠)	14(2个肉瘤) 15(2个癌) 56
8-甲基	0.4757	皮肤 皮下	20 20	12 9(1个癌)
6-甲基	0.4610	皮肤 皮下	20 20 60(大鼠)	11 11 58
11-甲基	0.4608	皮肤 皮下	20 20	2 6(1个癌)
5-甲基	0.4608	皮肤 皮下	20 20 60(大鼠)	0 2 4
4-甲基	0.3948	皮肤 皮下	20 20 60(大鼠)	0 0 3
10-甲基	0.3693	皮肤 皮下	20 20	1 0
2-甲基	0.3592	皮肤 皮下	20 20	0 1(癌)
9-甲基	0.3541	皮肤 皮下	20 20	1(肉瘤) 1
1-甲基	0.3540	皮肤 皮下	20 20 60(大鼠)	0 1 2
3-甲基	0.3185	皮肤 皮下	20 20	0 1

表 12-13

苯并[a]蒽甲基取代衍生物	作者计算的$\sum S_r^E-\sum S_r^N$值	Memory计算的K区S_r^E值	致癌性
单甲基衍生物			
7-甲基	0.6436	2.126	++++
12-甲基	0.5830	2.090	+++
8-甲基	0.4757	2.091	++
6-甲基	0.4610	2.180	++
5-甲基	0.4607	2.180	++
11-甲基	0.4606	2.082	++
4-甲基	0.3948	2.106	0
10-甲基	0.3693	2.091	+
2-甲基	0.3592	2.107	0
9-甲基	0.3541	2.082	+
1-甲基	0.3540	2.082	0
3-甲基	0.3185	2.083	0
二甲基衍生物			
7,12-二甲基	1.284	2.141	++++
7,7-二甲基	1.004	2.230	++
8,9-二甲基	0.867	2.094	+++
9,10-二甲基	0.729	2.094	+
2,9-二甲基	0.713	2.110	0
3,10-二甲基	0.688	2.102	0

应用K区 S_r^E 值所得结果间的比较。其它一些 PAH 甲基取代衍生物的 $\sum S_r^E-S_r^N$ 值跟实验致癌性的关系见表 12-14。

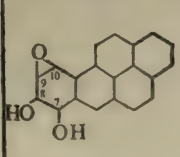
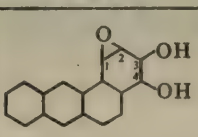
须要指出，湾区理论认为，在湾区形成的碳阳离子与 DNA 的结合是致癌的关键，因此湾区碳阳离子的稳定性与其母体烃的致癌活性应有平行的关系，亦即湾区碳阳离子的稳定性越高其致癌活性越强。用微扰分子轨道法计算也表明，强致癌的甲基取代 PAH 皆具有高的甲基阳离子离域能。实验证明，湾区致癌历程的关键步骤是 PAH 二醇环氧化物环氧背侧受到 DNA 亲核中心的进攻，生成一烷基化产物(参见 12.3.1)。我们用 EHMO 法计算的近似超离域度表明，苯并[a]芘的终代谢致癌活化分子——7,8-二醇-9,10-环氧苯并[a]芘与核酸碱基发生反应时，结合位置在苯并[a]芘的第 10 位碳，即湾区碳上；苯并[a]蒽的终代谢致癌活化分子之一的 3,4-二醇-1,2-环氧苯并[a]蒽与核酸碱基反应时的结合位置在苯并[a]蒽的第 1 位碳，也是湾区碳原子。计算结果与 Weinstein 等[43]的工作比较接近(表 12-15)。

然而，无论是湾区理论还是K区理论，对于许多复杂的 PAH 甲基取代衍生物结构致癌活性关系的处理和解释是同样不适用的。例如，图 12-32 中，单甲基二苯并芘(A和B)的致癌活性显著降低，而苯并芘本身的类似取代物(C)的活性却没有受到明显的干扰；在二甲基取代衍生物D和E中致癌活性消失，相反地，伴随F的去甲基化作用其致癌作用又得到恢复。相比之下，“双区理论”却有着明显的优点，它不仅可以对 PAH的结构致癌活性进行定性或定量处理，而且已成功地推广于各种单甲基多甲基取代的 PAH 上以及胆蒽类和萤蒽类等其它几类化学致癌剂的定

表 12-14

甲基取代衍生物	$\sum S_r' - \sum S_r''$	致癌性 皮肤	致癌性 皮下
苯并[a]芘			
3-甲基苯并[a]芘	0.5980		+++
12-甲基苯并[a]芘	0.5427		+++
4-甲基苯并[a]芘	0.5152		+++
7-甲基苯并[a]芘	0.5150		++
9-甲基苯并[a]芘	0.4320	−	−
3,6-二甲基苯并[a]芘	1.4812		+++
1,4-二甲基苯并[a]芘	1.1407		+++
4,5-二甲基苯并[a]芘	1.0896		+++
1,2-二甲基苯并[a]芘	0.9419		+++
2,3-二甲基苯并[a]芘	0.9060		+++
二苯并[ah]芘			
7-甲基二苯并[ah]芘	1.0710		++
3-甲基二苯并[ah]芘	0.4816		
5-甲基二苯并[ah]芘	0.4698		
䓛			
5-甲基䓛	0.3764	+++	++
11-甲基䓛	0.3764	+++	++
4-甲基䓛	0.3764		
10-甲基䓛	0.3761		
6-甲基䓛	0.3580		
9-甲基䓛	0.3451	−	
2-甲基䓛	0.3140	+	
1-甲基䓛	0.3073	−	−
5,6-二甲基䓛	0.8743	+	+
4,5-二甲基䓛	0.7535		+
2,3-二甲基䓛	0.6687		+
苯并[c]菲			
5-甲基苯并[c]菲	0.4112	+++	+++
4-甲基苯并[c]菲	0.3957	+++	++
3-甲基苯并[c]菲	0.3357	++	+

表 12-15 EHMO 法计算的苯并[a]芘和苯并[a]蒽，终代谢致癌活化分子的近似超离域度 S_r'

化合物	位置	S_r'	化合物	位置	S_r'
	7	0.0008		1	0.0120
	8	0.0005		2	0.0085
	9	0.0018		3	0.0075
	10	0.0217		4	0.0055

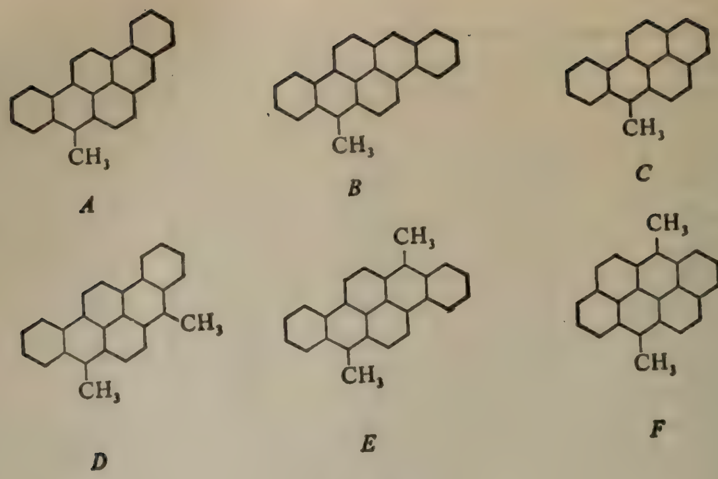


图 12-32

量关系上。

概括起来,我们可以看出,无论是K区、湾区或是双区理论对于PAH的终致癌形式是其环氧化物这点,在看法上是基本一致的,且都强调碳阳离子与生物大分子反应的重要性。我们知道,甲基通过电子效应能使碳阳离子稳定。研究表明,甲基化碳原子上发生的羟化反应,在活化历程中起着重要的作用,而甲基在生物体内也能代谢成亲电的甲基阳离子而参与多种生物化学反应。

总之,大量的事实已经说明,在一些位置上引进甲基基团后,可以加强PAH的致癌活性。而在另一些位置上引进后,致癌活性反而降底。量子生物学研究化学致癌作用的任务之一,即是进一步揭示甲基基团影响PAH致癌活性的机理,进而建立定量模型。

12.2.3 亚硝胺类

亚硝胺是一类广泛存在于自然界、人类食品和药物中的致癌物质。迄今已试验过的153种亚硝胺化合物,有118种致癌,占77%,而其中约有40种是环境、食品、药物中常见的。

自五十年代开始的对亚硝胺致癌活性的研究,揭示了亚硝胺在体内的代谢过程,良性肿瘤向恶性肿瘤发展的病理,可能的致癌机理等等。近十年来,有关亚硝胺的研究成果激增,研究内容也十分广泛,包括亚硝胺的体内代谢过程、毒性和毒理、细菌和动物实验数据、自然界和人体内的生成条件、分析方法、生化反应及其动力学、环境中的分布、各类食品和药物中的含量、与生物大分子的结合方式、致癌机理等等,以及用分子轨道法和"模式识别"法(pattern recognition method)对其分子结构与致癌活性关系的研究。

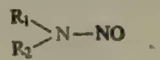
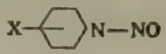
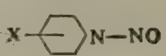
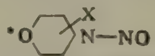
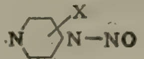
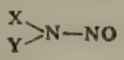
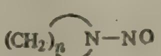
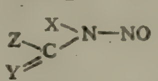
亚硝胺的通式为 $\begin{matrix} R_1 \diagdown \\ R_2 \diagup \end{matrix} N—NO$,当 $R_1=R_2=$烷基时,则称为亚硝胺;当 R_1(或 R_2)$=$酰基时,称为亚硝酰胺。通常所说的亚硝胺为这两类化合物的总称。

我们知道,分子轨道法是研究化合物结构的有效方法,但若应用于亚硝胺类化合物,则一般

只限于 R_1、R_2 较简单的情况，因为简单分子轨道法(如 HMO、EHMO 和 PMO 法)对于处理含共轭 π 键的平面分子尚属有效，而亚硝胺类化合物的分子系非平面分子；复杂的分子轨道法(如 SCF-LCAO-MO、PPP、CNDO 法等等)在列出亚硝胺类化合物的 Slater 行列式或者 FOCK 矩阵后，由于无法对其进行有效的“净化”，因此，即便用大容量的电子计算机求解，仍旧十分困难，况且计算结果的精度与“净化”程度有关，往往因为过分“净化”而引起较大误差。

目前国际上研究亚硝胺的结构与生物活性关系比较普遍的方法是采用“模式识别”(参见12.4)。该方法对研究对象首先进行结构分类，以决定引入参数的类型和数量，例如对 N-亚硝基类化物分成八大类分别进行计算(表 12-16)。我们在研究亚硝胺的结构与致癌活性关系时曾采用了这种分类方法。

表 12-16　N-硝基化合物致癌数据分布表

编号	结构类型	结构式	化合物数量				备注
			致癌(+)	弱致癌(W)	不致癌(−)	有争议(?)	
1	二烷基、二芳基(或混合)		47	6	18	3	
2	哌啶		10	2	7	6	
3	吡咯烷		3	1	4		
4	吗啉		2	1	1	0	*O或S
5	哌嗪		6		3	1	
6	肼羟胺		3		1	1	Y=O,N
7	杂氮环烷		5	1	0		n=3～7
8	酰基化物		29	2	1	7	Y=O,N Z=C、N、O
小计			105	13	35	18	

我们使用经验电子参数，仅引入两个参量对第一类直链亚硝胺化合物进行了计算，并得到了一个简单的初等函数作为数学模型的参考式，其验证计算值与实验活性值的符合率为 97%。[44]

12.2.4　微生物毒素

在这一节里，我们主要讨论致癌性霉菌毒素。

致癌性霉菌毒素是曲霉分泌的一种致癌物质。曲霉属系真菌，分布很广，从空气、土壤、水、直到人类所接触的各种环境中都广泛存在，它与人类生活密切相关。

对于致癌性霉菌毒素的研究是六十年代后才开始的。1960 年英国等国发现黄曲霉素能强烈地引致动物肝癌。随后又陆续发现了杂色曲霉毒素、黄米霉素、环氯肽、棒青霉素、灰黄霉素等。其中，黄曲霉素 B_1(AFB_1)(图 12-33)，是一种强力的致癌剂和诱发肿瘤的霉菌毒素，对雏鸭 $LD_{50}=0.5\ mg/kg$。以 400 μg 微量的黄曲霉素 B_1 投与实验雄鼠，经 68 周后，100%均发生肝癌。流行病学考查证明，肝癌的发病率与黄曲霉素的污染有着平行关系。此外，还能引起食管、胃、结肠等部位的癌症。

对于致癌性霉菌毒素的致癌机理的研究还刚开始不久。近年来研究了黄曲霉素的代谢过程，并且分离出黄曲霉素与腺嘌呤的共价结合产物，并证明其 2,3-双键的环氧化物是其最终致癌形式。若 2,3-双键被 Cl_2 加成，其二氯化物毋需活化就是强致癌剂。

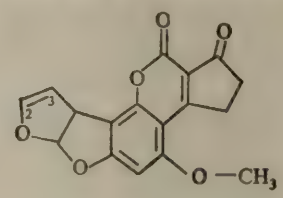
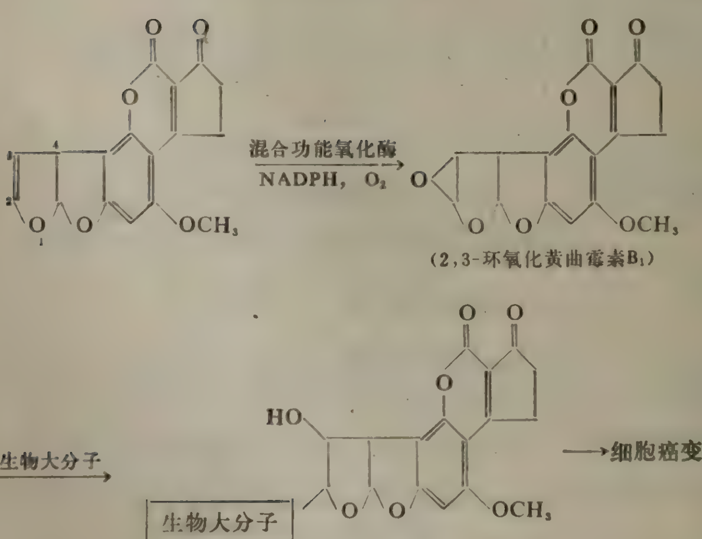


图 12-33　黄曲霉素 B_1



黄曲霉素 B_1 代谢后转化为 2,3-环氧化物，它可以与核酸或单、多核苷酸形成加成物[45,46]。与核酸的加成发生在 AFB_1 的 C 2 和鸟嘌呤平面的 N 7 之间。这个结构与从 AFB_1-2,3 氧化物在 C 2 位置形成碳阳离子是一致的。研究还表明，加成物形成的程度为：DNA 比 RNA 或最活泼的聚(G)要大；与单核苷酸的反应性远远小于与多核苷酸的反应性；与 GMP、AMP、CMP 的反应性比 UMP 或 dTMPT。这些结果与单个碱基、核苷、核苷酸、单股螺旋和双螺旋分子静电势的计算是一致的(参见 4.8)。

12.2.5 多环芳烃“双区”理论的发展

自 1979 年戴乾圜建立以离域能为参数的“双区”理论以来，又以双区理论为基础，开展了更系统的研究工作，取得了较大的进展。以下我们将从三个方面加以简介。

12.2.5.1 多环芳烃“双区”理论的定量研究

陈洪等[47]在应用概率统计方法和回归分析的一般原则，验证“双区”理论及其定量公式合理性的基础上，将距离参数(D)，传输参数(C)和结构参数(L'/L)引入多环芳烃结构致癌活性关系的定量公式(12.1)式之中，从而提出了双区公式的其它新形式(12.8 和 12.9 式)：

$$\log K=[6^2-(D-2.5)^2]^{0.4}\left(\frac{L'}{L}\right)^{0.3}\Delta E_1\cdot\Delta E_2-\left(\frac{C}{L}\right)n\cdot\Delta E_2^3 \tag{12.8}$$

$$\log K=1.1-\left(\frac{D}{1.39}-2\right)^2\cdot1.1-(C-5)^2\left[\left(\frac{L'}{L}\right)^{1/6}\lambda_1-0.1\,n\lambda_2\right] \tag{12.9}$$

对于(12.10)式规定： $0\leqslant K\leqslant11$ 为“－”；$11<K\leqslant15$ 为“＋”； $15<K\leqslant30$ 为“++”；$30<K\leqslant50$ 为“+++”；$K>50$ 为“++++”。公式(12.9)在整个规定的指标内，置信度都超过 95%。计算结果与实验基本一致。

式(12.9)的特点是：(1)等式两边均为常数，不存在量纲不统一的问题，为今后的理论探索提供了方便；(2)能反映出每个参量对函数的独立贡献，每一项都有独立的物理意义；(3)用(12.9)式计算了目前已作过生物实验的 63 个 PAH，都得出了比较满意的结果，尤其是对于八稠环以上的 PAH。

陈洪等采取了能量商的概念，能量商愈大则可能增加反应发生在特定部位的几率，因此与采取能量的绝对值相比，能更好地模拟反应的几率。λ_1、λ_2 的引入就是企图反映这个规律，而能量商最大的部位可能在K区(即活性K区)、角环、次角环等等，双区公式(12.1)式是根据超热力学假定得到的，因此使用 ΔE_1，ΔE_2 是合理的。式(12.9)是根据双区理论应用统计分析得到的，我们知道在作回归分析时，应该消除参量所取单位而造成的影响，即进行所谓“标准化”处理以消除量纲(参见 12.4.2)。因此，这时以取 λ_1、λ_2 比用 ΔE_1，ΔE_2 更合理些。

他们认为，PAH 经代谢活化后与生物大分子作用的第一优先区域从统计上看应为能量商最大的区域(K区，角环，甚至L区中的一个碳原子)。第二优先区域，在统计上应为空间因素有利的角环或次角环。这种选定区域方式不致产生含混，可能在统计上更符合生物代谢的模型。(12.9)式，对于八环的计算，根据现有的数据是符合的，且较(12.1)式可能有更广泛的适用性。但是，我们也不难看出，(12.1)式较(12.9)式更简捷，对八环以下的 PAH 的计算较为方便。

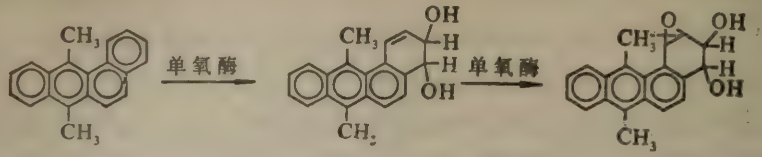
12.2.5.2 甲基、烷基代多环芳烃等的定量结构致癌活性关系的研究

在 12.2.2 中，我们已经简述过甲基取代 PAH 的结构与致癌活性的问题。

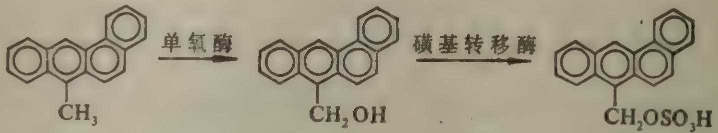
鉴于戴乾圜等运用双区理论首次合理而系统地建立了 187 个甲基、烷基、环戊并、环己并 PAH 以及胆蒽系的定量结构致癌活性关系，并使计算与实验的符合率达 95%以上。为此，我们认为有必要在本书的这一章里，反映他们在这一研究领域中所获得的一些进展。

首先，关于烷基代多环芳烃(APAH)的所谓“终致癌剂”问题，目前存在着两种对立的观点。

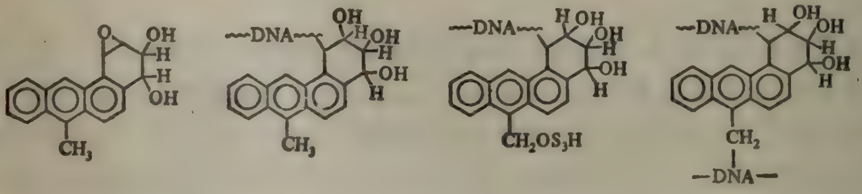
一种观点认为，APAH 湾区角环的四氢二醇环氧化物是所谓的“终致癌剂”。例如近年的实验证明，7,12-二甲苯并[a]蒽有如下代谢：



另一种对立的观点则认为，烷基氧化后生成的 α-醇酯，尤其是磺酸酯，应是 APAH 的所谓“终致癌剂”。例如也有许多实验支持如下的代谢：



上述实验事实，用 DNA 的单股烷化观点是无法加以统一的。但是，若从双官能烷化观点（双区理论实际上是一种双官能亲电理论）出发，则上述实验事实皆可以得到合理的阐明。在 DNA 股间双官能烷化的重要步骤之一，很可能是经角环 E 碳的环氧化物，再进一步经 α-甲磺酸酯而发生横向交联：



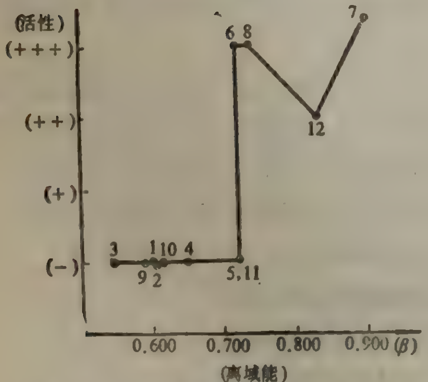


图 12-34　甲基代苯并[a]蒽中 α-碳离域能与致癌活性关系（图中数字为甲基位置）

当然，这决非唯一的步骤。显然可能在代谢活化的各个亲电中心之间，还可能发生其它形式的交联。可能在角环，次角环，活性 K 区以及甲基之间发生各种组合各种顺序的横向交联，虽然其趋势可能较上述趋势为弱。

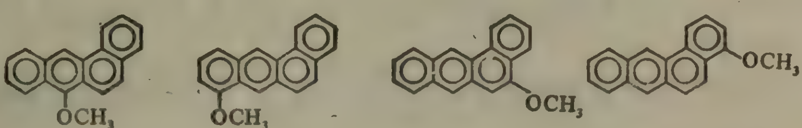
双区理论使得 APAH 结构活性关系的种种实验事实，皆得到合理的解释。根据 PMO 计算及 12 个甲基代苯并[a]蒽的综合概括实验数据，可大致将其致癌活性与亚甲基阳离子离域能数值绘成图 12-34。可见致癌活性随亚甲基的亲电烷化趋势的增加而增加。角环 E 碳的亲电烷化趋势，应该是大致类似的。因此，上述趋势实际上反映着分子经代谢后的横向交联能力与其致癌性的相关关系。总

之，甲基 PAH 的代谢过程中，角环湾区阳离子及亚甲基阳离子孰是“终致癌剂”的有争议的问题，按照双区理论可以取得统一。

再者，甲基对于 PAH 的代谢活化，可能产生两类空间影响。一类增强 PAH 的致癌活性，可称之为增活位阻。另一类则减弱 PAH 的活性，称之为减活位阻。

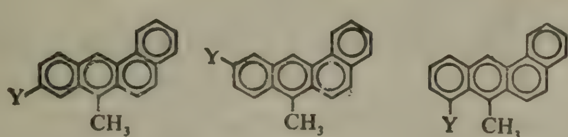
甲基对芳环有诱导效应，场效应和动态超共轭效应，从而影响环上有关位置的反应活性。甲基在基态及静态时不存在超共轭效应，但在离子、自由基等瞬变体中，在激发态或进攻试剂影响下，可动态地产生超共轭效应。

我们假定甲基起着动态的斥电子共轭效应（实际是甲基的被极化）而对环上特定的位置发生有利于代谢的影响。根据相应取代 PAH 的等共轭亚甲基 PAH 离子的离域能(PMO)，可以探讨于下，如：



0.879(++)　　0.737(++)　　0.722(+)　　0.647(-)

甲氧基是难于代谢水解的，离域能大的前三者，由于甲氧基斥电子使电子更有效地转移到角环、次角环、K 区等活性区域上。这利于活化而使其从近乎不致癌的母体向中等致癌转化。自然，后二者尚有减活位阻的影响。相反，拖电子基则表现出对 7-甲苯并[a]蒽(+++)致癌活性降低的影响。



Y= CN(-) / Cl(-) 0.593　　Y= CN(++) / Cl(+) 0.614　　Y= CN(+++) / Cl(+++) 0.737

离域能愈大，则意味着 CN 和 Cl 从母核上夺取的电子较少，而对 PAH 的活性影响较小。

再就是对烷基代多环芳烃的定量处理所作的研究。双区理论的公式是根据 Gibbs 关系式以及 PAH 活化与脱毒的代谢过程，对理想状态的细胞，经数值拟合而得的半经验公式[(12.1)式]。

甲基的各种影响，可用适当的参数修正 ΔE 而定量地体现，所有的增活位阻，用定域键序 1 除以可能最大的离域键序 0.8 乘以 ΔE_1 来校正。对于减活位阻，由于此类位阻的存在可能更重要的是影响第二个活性区的代谢与交联，因此，ΔE_1 可近似地毋需校正，而在 ΔE_2 上减去 0.12 β。存在两个以上的减活位阻则减去 0.24 β。

脱毒区的脱毒趋势与有关位置碳原子的定域能（L 区、等价 K 区、角环）或甲基正离子的离域能有关。对 PAH 母体我们假定脱毒区的脱毒趋势彼此等值，可以得到良好的结果。而烷基代 PAH 中，甲基、烷基等的空间效应极大地影响着脱毒区的脱毒代谢。同时甲基等脱毒区的脱毒与活化的主导趋势更为敏感地取决于离域能。因此，对多于两个亲电中心要求的甲基按表 12-17

规定选取 n 值。由于角环及次角环 M 碳及 M′ 碳的定域能与相应离子的离域能有相关性。为方便计，也以离域能统一地判别取值。如表 12-17。

表 12-17 脱毒区数目因位阻而减少的取值规则

ΔE(β)	无位阻	有位阻	有位阻(乙基以上烷基)	备　注
ΔE<0.6	$n=2$	$n=1$	$n=0$	适用多余的甲基、次角环，角环
0.6≤ΔE<0.8	$n=1$	$n=0$	$n=0$	
ΔE≥0.8	$n=0$	$n=0$	$n=0$	
L区，等价K区	$n=1$	$n=0$	$n=0$	

戴乾圜等还根据一些定量参数[19]，运用双区公式(12.1)式，对 187 个烷基代 PAH，计算与实验结果相符者达 95%以上。

从总体上看，烷基代 PAH 定量处理的成功，使我们看到，双区理论的某些合理性和成功之处。

12.2.5.3 多环芳烃的致癌性与其分配系数之间的定量关系

戴乾圜的双区理论定量公式在对超过两百多个 PAH 致癌性的计算中获得了较大的成功[6,19,48]，偏差很小。然而，由于双区定量公式(12.1)式忽略了传输因素等，因此不适用于八稠环以上的 PAH。为了解决对八稠环以上的 PAH 的致癌性计算所出现的偏差，陈洪与戴乾圜又提出了新的双区定量公式(12.9)式。式(12.9) 尽管比较成功，但公式比较复杂，不易推广，因此存在着很大的改进余地。

近来，王连生等[49]以双区理论为基础，提出了 PAH 的致癌活性与其分配系数之间的定量关系方程：

$$\log K = 5.30\Delta E_1 \cdot \Delta E_2^3 - 0.15(1.3\,n + |\log K_{ow} - 6.401|) \qquad (12.10)$$

式中的 ΔE_1、ΔE_2 分别为双亲电中心的离域能，n 为脱毒区总数，K 为致癌性指标。

经验算，王连生等选用最合适的分配系数为

$$\log K_{ow} = 6.40$$

他们利用式(12.10)计算了 77 个 PAH 的致癌性，计算值与实验值基本符合。对八稠环以上的 PAH 的计算也获得了较好的效果。

综上所述，在双区理论的基础上计算致癌活性的双区定量公式存在着很大的改进和应用前景，而不断的改进将使双区理论更加臻于完善，应用范围将不断扩大。

12.3 化学致癌物同生物大分子的相互作用

化学致癌物同生物大分子(DNA、RNA 和蛋白质)的相互作用，长期以来一直受到化学致癌研究者的重视。大量的实验研究显示，许多化学致癌物在体内经代谢活化，而形成活泼的亲电性物质，它们能以共价键与细胞的核酸和蛋白质结合。

12.3.1 同核酸的相互作用

DNA 的 A、G、C 和 T 四种碱基都能同化学致癌物结合，有时磷酸二酯骨架也受到致癌物的影响。从化学致癌物与碱基结合的位置来看，鸟嘌呤的 N 7 和N 2 位置似乎最为活泼，其次是腺嘌呤的 N 3 和 N 6 位置。RNA 同大多数致癌物结合的范围和 DNA 类似，甚至更多样化些。表 12-18 列出了一些化学致癌物与 DNA 的结合。

从表 12-18 可以看出，不同的致癌物可以与 DNA 分子中不同位置结合。其中，直接致癌物——烷化剂通常是与鸟嘌呤的 N7或 O6结合。O6的取代和N7的取代不同，按 Watson-Crick 的双螺旋结构模型，O 6 的取代能够直接影响碱基配对，而且 O 6 取代后在敏感组织的修复也较 N7 取代更为缓慢。N 7 的烷化能引起嘌呤丢失，从而减少了 DNA 修复机制的忠实性。表 12-19 为双链核酸在体外的烷基化位置和烷基化百分率，然而芳香胺类致癌物则与鸟嘌呤 C 8 结合，而有些致癌物(如杂类)则往往与鸟嘌呤的 O 6 结合。在 DNA 的四种碱基中，胞嘧啶、腺嘌呤和胸腺嘧啶各个位置的原子则较少受到化学致癌物的攻击。

表 12-18 致癌物与DNA的结合

致癌物	取代或反应的代谢形式	改变部位				效果
		G	A	C	T	
烷化剂						
DMN	CH_3—,C_2H_5—等	N7	N3	N3	O4	去嘌呤化
MNU		O6	N1	N1	N3	X线型修复
MNNG		N3	N7			碱基取代
MMS等		还有磷酸基团				突变
芳香胺						碱基移位
AAF	═C(—)—N—O—SO_3	C8	?	?	?	紫外线型修复
MAB?		N2				框移突变
多环芳烃						
DMBA	DMBA-5,6-氧化物	N2	N6	N4		碱基移位(?)
BP	B[a]P-7,8-二氢二醇-9,10-环氧化物	N2	N6	N4	C6?	紫外线型修复
7-BrDMBA	7-甲基-DMBA	N2	N6	N4		框移突变
杂类						
黄樟素	1′-乙酰氧基	O6				
黄曲霉毒素	-2,3-氧化物	?	?	?	?	
4-NQO	-4-OH-氨基	?	?	?	?	
金属	?	?	?	?	?	
有机氯	?	?	?	?	?	

现在，我们以苯并[a]芘最终致癌物——7,8,9,10-四氢-7,8-二羟基-9,10-环氧衍生物为例讨论它和核酸间的结合方式。

根据实验结果，可以认为，苯并[a]芘的最终致癌物在与 DNA 及 RNA 发生共价结合时，主

表 12-19 双链核酸在体外的烷基化*

	总烷基化百分率					
	Me_2SO_4	MeMS	MeNU	Et_2SO_4	EtMS	EtNU
腺嘌呤						
N1	1.9	3.8	1.3	2	1.7	0.2
N3	18	10.4	9	10	4.9	4.0
N7	1.9	(1.8)	1.7	1.5	1.1	0.3
鸟嘌呤						
N3	1.1	(0.6)	0.8	0.9	0.9	0.6
O6	0.2	(0.3)	6.3	0.2	2	7.8
N7	74	83	67	67	65	11.5
胸腺嘧啶						
O2			0.11	nd	nd	7.4
N3			0.3	nd	nd	0.8
O4			0.4	nd	nd	2.5
胞嘧啶						
O2	(nd)	(nd)	0.1	nd	nd	3.5
N3	(<2)	(<1)	0.6	0.7	0.6	0.2

* nd 表示衍生物不是检出的
Me_2SO_4——硫酸二甲酯
MeMS——甲基甲硫磺酸酯
MeNU——甲基亚硝基脲
Et_2SO_4——硫酸二乙酯
EtMS——乙基甲硫磺酸酯
EtNU——乙基亚硝基脲

要是与鸟嘌呤结合。其结合的位置发生在苯并[a]芘最终致癌物的第 10 位碳(C 10)与鸟嘌呤核苷酸的第 2 位氨基基团之间(图 12-35*A*),也可在苯并[a]芘最终致癌物的第 10 位碳与鸟嘌呤核苷酸的 N 7 位置(图 12-35*B*)或腺嘌呤的 N 6 位之间结合(图 12-35 *C*)。

Pullman 等[50,51]曾应用 ab initio SCF 法对苯并[a]芘代谢致癌物跟碱基(G、A 和 C)的结合模型作了量子力学计算。他们具体处理了苯并[a]芘代谢致癌物与B-DNA中鸟嘌呤、腺嘌呤和胞嘧啶的氨基的结合物(参见图 12-36、图12-37),并得到了相应的 BPTC-核苷中间体和 BPTC-核苷结合物的构象图(图 12-37 表示后者)。

Pullman 等将苯并[a]芘二醇环氧化物跟核酸碱基的氨基间形成结合物的过程,分作两步进行计算(见图 12-36)。最终结合物和中间物的构象系用经典势计算的。计算表明,如果苯并[a]芘的代谢环取 B(I)(船式)或 HC(I)(半椅式)构象,则结合物和中间物呈现较小的空间位阻(参见图 12-38、表 12-20)。

在 BPTC-核苷结合物非键相互作用能的总体极小值的计算结果中,以 BPTC-脱氧鸟苷 B(I)构象体的能量最低。

从 Pullman 等的计算结果可以看到,BPDE 跟 DNA 结合,主要是其中的鸟嘌呤的氨基,这跟实验结果是符合的(参见表 12-18)。此外,还可以看出,苯并[a]芘代谢物与 DNA 的鸟嘌呤、

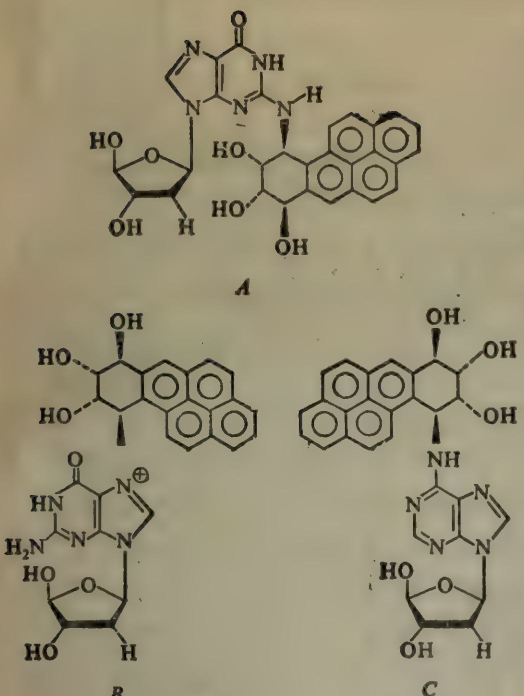


图 12-35

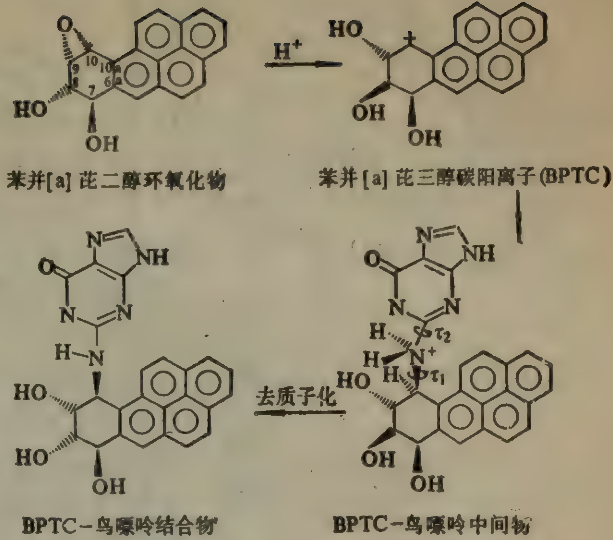


图 12-36

表 12-20 BPTC-核苷结合物非键相互作用能的总体极小值(kcal/mol)*

核苷 \ 环构象体	B(I)	HC(I)	B(II)	HC(II)
脱氧鸟苷	0.58 (−0.02)	2.16	3.95	3.22
脱氧腺苷	2.96 (2.24)	4.30	5.75	5.93
脱氧胞苷	2.83	3.78	4.96	5.25

* 括号内的数值相应于 $rN_{amino-C}=1.43$ Å

腺嘌呤和胞嘧啶的相互作用,其作用快慢顺序为:鸟嘌呤＞腺嘌呤＞胞嘧啶。从分子静电势的计算值(参见 4.8)来看,在 DNA 双螺旋中,静电势深度的次序为鸟嘌呤＞腺嘌呤＞胞嘧啶。这样的次序与碱基与 PAH 的湾区代谢物的亲合力有关。

PAH 与细胞中 DNA 结合的数量,据研究[52]大约为每 $2\sim3\times10^5$ 个核苷酸中有一个与致癌物相结合。这个量从化学角度来看是微不足道的,但从生物学角度来看意义就不一样了。因为一个细胞约有 1×10^9 对碱基,这就意味着,每个细胞中与致癌物共价结合的碱基数约为 1×10^4 个。这已经是一个不算小的数目了。在致癌物的长期作用下,一旦 DNA 修复机制失灵,就必然会引起 DNA 的损伤,并由此造成一系列转录和翻译过程的失调。

一般认为,除了化学致癌物同核酸之间的共价结合外,至少还存在着两种非共价结合的情

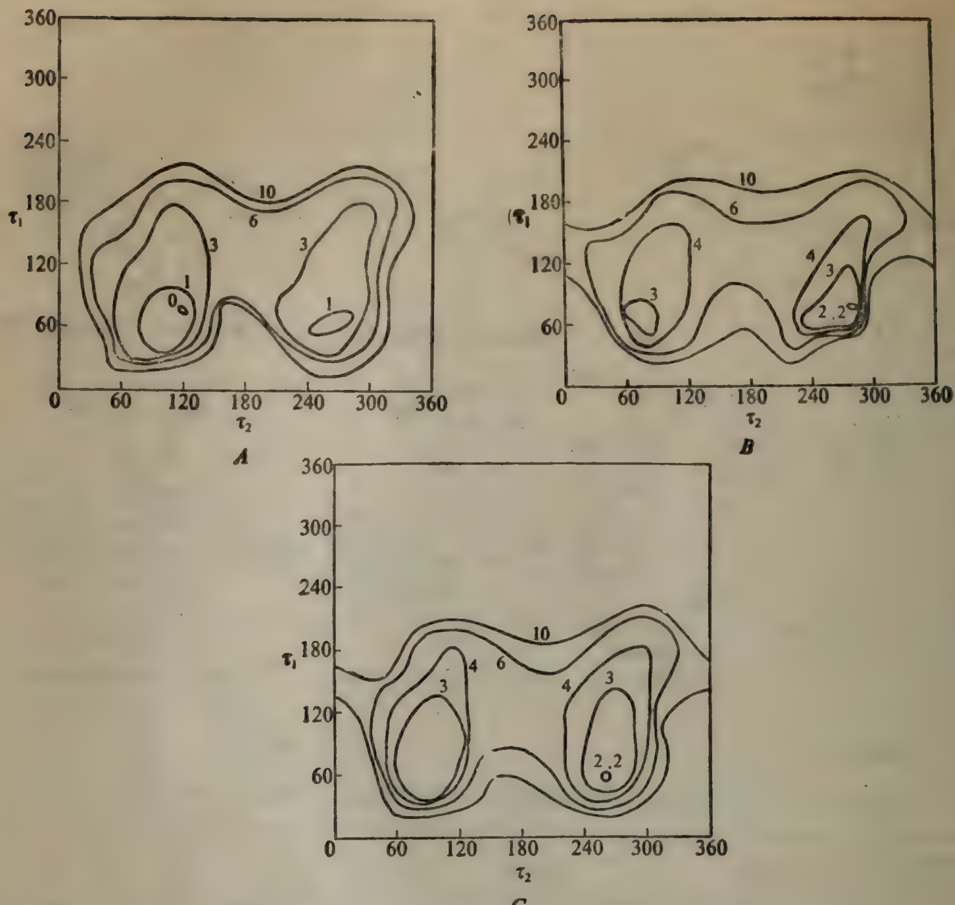


图 12-37 BPTC-鸟嘌呤核苷(*A*)、腺嘌呤核苷(*B*)和胞嘧啶核苷(*C*)结合物的构象图
(苯环构象体为 B(I))

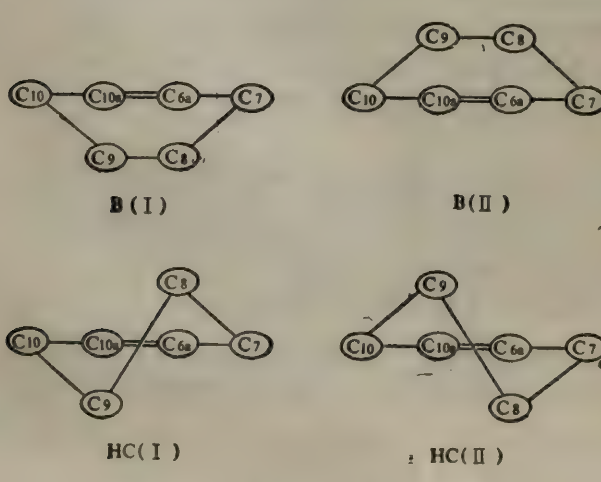
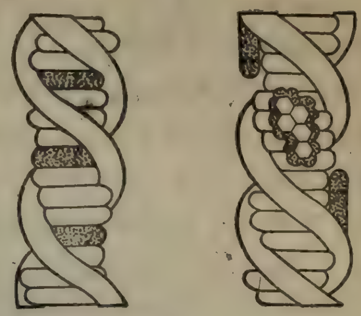


图 12-38 苯并[a]芘代谢苯环的构象体

况。一种为嵌入，它是指在 DNA 的碱基对之间插入或嵌入化学致癌物分子(图 12-39*A*)。当这类化合物嵌入 DNA 碱基对之间以后，DNA 双螺旋便有可能部分地被解开，从而影响到 DNA 执行其正常的功能。第二种更弱的非共价结合，是致癌物分子塞进 DNA 螺旋。在此情况下，致癌物分子成轴状结合，垂直于 DNA 的碱基(图 12-39*B*)，这种结合方式已经得到证实[53]，其致癌作用的意义目前尚不清楚。

然而，若干研究结果说明，化学致癌的初始阶段，其关键因素是核酸-化学致癌物结合物的生成。而这些结合物的化学结构的阐明，对于进一步了解化学致癌作用无疑是极其重要的。最近，Stezowski 等人[54]的工作注意到了活性致癌物使 DNA 核苷酸烷基化后对后者立体化学性质的影响。我们认为，在探讨化学致癌物同核酸的相互作用时，必须充分注意到核酸等生物大分子与致癌物相互作用的立体化学以及 B-DNA 和 Z-DNA 与致癌物结合的情形。



A　　*B*

图 12-39　化学致癌物同 DNA 之间的两种非共价结合：*A* 表示嵌入，*B*表示塞进DNA螺旋，与碱基垂直

近几年来，人们已逐渐发现，化学致癌物要通过对细胞内 DNA 等大分子的作用才能致癌，也就是说细胞内 DNA 的异常是癌产生的内在根据。究竟哪些因素影响着化学致癌物与核酸的亲电反应，这已经可以应用分子静电势的概念作出量子力学计算的解释[55]。

在第 4 章，我们已计算了核酸碱基上各位置在不同情况下的静电势最小值(参见 4.8)，并对其变化规律作过一般讨论。由于负的分子静电势表示对正电性的吸引，当考虑表中计算结果的实际应用时，我们首先想到的是那些与碱基有关的相互作用，特别令人感兴趣的是化学致癌物质。因为业已证明，许多化学致癌物(所谓前致癌物或最终致癌物)的活性形式是阳离子或带有缺电子原子反应中心的不带电分子。致癌物有的本身就是活性的(许多相对比较简单的烷化剂)，有的必须经过代谢活化才具有活性(如芳香烃、芳香胺、黄曲霉素 B_1 等等)。核酸的碱基是这些化学致癌物质的最重要的靶物质之一。由计算知道，碱基、核苷、核苷酸、单螺旋、双螺旋或变性 DNA 的环氮 N 和羰基氧原子 O 的分子静电势深度逐步增加，实验表明，跟这种变化相一致的是它们与亲电性致癌物质的反应性也随静电势深度的增大而增加。因此，我们的计算可为这些亲电性致癌物质和核酸及其组分的相互作用提供一个统一解释。

12.3.2　同蛋白质的相互作用

根据迄今为止的大量实验事实，研究报告和有关“化学致癌”的专著所报道的结果和所持的观点来看，在活细胞的生物大分子中，DNA是大多数化学致癌物攻击的最主要的靶分子，而RNA和蛋白质等其它生物大分子则仅仅处于被致癌物攻击的次要地位，甚至处于辅助的地位。在结构致癌活性关系的研究中，有人假定 DNA 分子是致癌作用的唯一靶物，并认为只有 DNA 分子而不是 RNA 或蛋白质，可以长期储存致癌物所诱发的信息[12]。从理论上讲，核酸和蛋白质都能通过亲电、亲核基或协同反应与化学致癌物共价结合。然而，绝大多数人则认为，化学致癌

作用的靶分子是 DNA。其实,当人们注意到致癌物与核酸结合的某些实验结果和因结合物的形成而引起的生物学功能改变之前,就已经知道了致癌物与蛋白质的共价结合。但是,有关蛋白质与致癌物结合的位置、结构和由此而引起的生物学功能的改变,至今都还知道得很少。出现这些情况的原因,除了 DNA 可能在事实上确实是致癌物的最主要靶分子外,跟生物化学研究中,那种至今还比较普遍存在的"重核酸轻蛋白质"的现象,跟 DNA 双螺旋结构模型建立以来的分子生物学、分子遗传学等等所取得的巨大成就是分不开的。在这当中,DNA 双螺旋、信使、中心法则和遗传密码等新概念的引入和发展起了极大的作用。

尽管目前我们对致癌物同蛋白质结合的问题还缺乏更多的了解,但是,已经知道某些致癌物怎样与蛋白质以及与构成蛋白质的氨基酸发生反应。例如,芳胺或偶氮染料可能通过甲硫氨酸与蛋白质结合;而另一种致癌物二甲基亚硝胺则通过一个甲基基团引入组氨酸来与蛋白质发生反应。但是,另一种亚硝基化合物(N-甲基-N′-硝基-N-亚硝基胍)却具有一种特异的效应,一方面可以通过加上甲基而使 DNA 甲基化,另一方面则可以通过将—C(NH)—NH—NO_2 加到蛋白质的氨基酸(赖氨酸或半胱氨酸)上,而使蛋白质酰硝胺化。

近来,Kootstra 等[13,14]就组蛋白的乙酰化作用对苯并[a]芘同染色体结合物形成和迁移的影响以及苯并[a]芘最终致癌物同鼠皮肤的 DNA、RNA 和蛋白质的结合等进行了一系列的研究。此外,Sculley 等[15]对苯并[a]芘以及(±)-7β,8α-二羟基-9α,10α-环氧-7,8、9、10-四氢苯并[a]芘同组蛋白的相互作用进行了研究,结果显示出对于受致癌物的攻击,在组蛋白上相应的亲核性受体部位的存在可能是必要的,且致癌物同组蛋白的相互作用,在导致一系列转化过程中或许是一个重要的步骤。他们还看到,组蛋白 H_1、H_{2A} 和H_3 对于[^{3}H]-7,8-二羟基-9,10-环氧-7,8,9,10-四氢苯并[a]芘呈现强标记,反之,H_{2B}、H_4 的放射活性却不显著,且无论在体内或是在试管内,都看到鼠 10T1/2 细胞和人肺上皮细胞核心组蛋白 H_{2A} 和H_3因为同苯并[a]芘最终致癌物结合而发生改变。Macleod 等[16]的研究也证实,在体内实验中仓鼠胚胎细胞组蛋白 H_{2A} 和 H_3,因为同苯并[a]芘发生共价结合而受到修饰,并且也看到其它致癌性碳氢化合物同核心组蛋白的结合也有类似的特征。所有这些研究工作对于人们了解化学致癌物同蛋白质的相互作用,从一个侧面提供了一些有意义的启示。但距离真正揭示化学致癌物同蛋白质相互作用的意义,还有相当的路程要走。

12.3.3 致癌物作用后大分子的改变

无数的实验事实告诉我们,核酸和蛋白质等生物大分子同致癌物相互作用后,其结构、合成和功能都有可能因受致癌物的作用而发生改变。

从表 12-18 可以看出,PAH、黄樟素和一些烷化剂同鸟嘌呤的 O6 和 N2,腺嘌呤的N6,胸腺嘧啶的 O4 和胞嘧啶的 N4、N3 等结合,而这些都是碱基配对时发生氢键结合的位置。一些实验研究表明,双功能基的烷化剂能在 DNA 的双链间,或一个链的两个部位之间或两个 DNA 双链之间形成交叉连结,从而干扰DNA的复制和转录。此外,如前所述,一些致癌物分子嵌入在DNA碱基对之间以及致癌物塞进 DNA 双螺旋也将会影响到 DNA 的结构和功能。例如改变 DNA 双螺旋结构和构型,在电子显微镜下可以看到由于致癌物质引起的局部变性。

事实上，不同类型的化学致癌物同 DNA(或 RNA 和蛋白质)的结合，可以有多种形式的相互作用，已经研究了几种可能性：

——DNA 发生改变，

——RNA 发生改变，转而改变 DNA

——蛋白质发生改变，因此改变了 DNA 的表达。

但是，致癌物与生物大分子的结合及其致癌性之间，有的是相关的，有的是不相关的。这可能有的是由于对细胞没有产生重要的影响；有的细胞由于生物大分子受到影响而发生改变后，通过修复能使该细胞仍有补偿机能，有的则引起分子的连锁反应，影响到后代。后一现象特别引人注目。

12.4 化学致癌作用的量子生物学研究方法

研究实践告诉我们，要真正弄清化学致癌这样一个复杂的问题，并非是某一个单一理论或方法所能及的。为探索致癌之谜，各种不同学说和方法的提出是很自然的，其中各种方法的应用和补充在一定程度上推动了化学致癌作用研究的向前发展。仅就量子生物学研究方法来看，它实际上包含了量子化学方法、多元回归分析法和模式识别法等等方法。

12.4.1 量子化学方法

正如我们已经知道的那样，迄今量子生物学所用的理论计算方法基本上是沿袭量子化学的方法。因为量子化学方法可以计算分子中的电子，特别是价电子能级、分子的电荷分布、键级、自由价等结构特性。特别是为那些实验非常难于测定的反应中间体提供微观图式。

在化学致癌作用的研究中，大致有如下应用：

(1) 结构与活性

这类工作常在系列化合物的致癌、诱变、毒性、降解性、生物积累和量子化学参数间建立关联，探讨分子官能团所起的作用等等。白迺彬曾用 HMO 法，辅以模式识别技术探讨了 PAH 结构与其致癌能力的关系。

(2) 立体结构

有些化学致癌物的化学特性因本身的构型、构象不同而截然不同。量子化学则是从分子结构角度研究构型、构象引起特性差异的有力工具。本章作者白迺彬等曾用CNDO法研究了 PAH 最终致癌物构型、构象差异之起因(参见 12.2.1.5)。

(3) 反应途径

化学致癌物在体内的代谢转化往往涉及若干反应历程和自由基，形成许多中间产物，有许多反应方向及其终结途径。然而，我们知道，反应途径往往相当难于进行实验研究。但是，量子化学计算则有助于辅助实验判断反应途径。

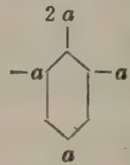
有关量子化学中的分子轨道法在化学致癌研究中的应用，特别是一些半经验方法的应用，我们已经作过较详尽的讨论，下面我们将补充介绍孤立分子近似和位能面等，以利读者了解化学致

癌研究中量子化学方法的概貌。

12.4.1.1 孤立分子近似[56]

孤立分子近似把致癌剂对靶分子(核酸和蛋白质分子等)的作用看成是对靶分子的微扰，它使用靶分子(受体分子)在基态时的一些分子轨道参量来指示靶分子的反应活性。我们知道，许多致癌物(特别是 PAH)的代谢致癌过程,包含着阳离子(亲电试剂)或阴离子(亲核试剂)对受体的进攻。根据有机反应的电子理论，这些离子试剂总是优先攻击那些带有最大相反电荷的原子。在 π 电子的分子轨道理论中,是用 π 电荷密度来描述分子中的电子分布。因此我们预言,在亲电反应中,电荷密度最大的位置优先反应;亲核反应中,电荷密度最小的位置首先受到攻击(实际上,这里的电荷密度应该包括 σ 电子的贡献,不过在 π 电子理论中，总是忽略 σ 电荷密度的任何变动)。这个相当朴实的概念,颇能成功地说明反应活性。

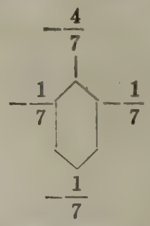
从简单分子轨道理论知,凡是与奇交替烃离子等共轭的体系,它的 HMO 电荷密度可以用非键分子轨道系数来计算。例如,重氮盐和苯酚的偶联反应以重氮离子 RN_2^+ 为进攻试剂,$C_6H_5O^-$ 为受体。其中 $C_6H_5O^-$与苄基负离子 $C_6H_5CH_2^-$等共轭。苄基的非键分子轨道系数是



$$a^2+a^2+a^2+4a^2=1,$$

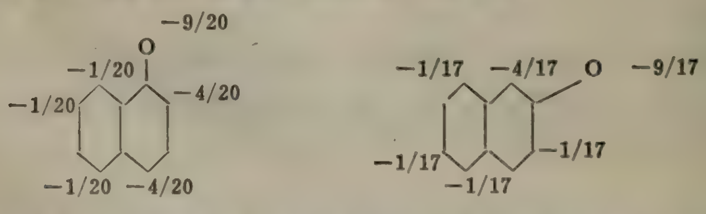
$$a=1/\sqrt{7}$$

苄基负离子的负电荷按照非键分子轨道系数的平方来分配，即



由此预言苯酚的偶联反应将发生在羟基的邻位或对位，这和实际符合。

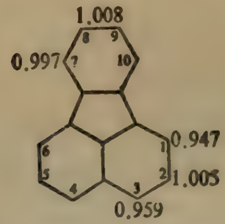
用同样的方法计算 1-萘酚和 2-萘酚的电荷分布,其结果是



因此 1-萘酚是 2 位和 4 位,2-萘酚是 1 位发生偶联反应。如果 2-萘酚的 1 位已被取代，便很难在其它位置发生偶联反应。这些也都符合实验。

当处理非交替烃萤蒽时,HMO 法的电荷分布(如下页图所示)。

按此预言活性顺序是 8>2>7>3>1,与实测的亲电取代顺序(例如 50°C时在醋酸中硝化)3>8>7>1>2不符。一般地,孤立分子近似对杂环系或非交替烃的预言往往不太可靠。



用电荷密度指示反应性能只限于电荷分布不均匀的分子。对于中性交替烃，它的 q_r 处是 1，此时可以改用原子自极化率 $\pi_{r,r}$ 来作反应活性指标。原子自极化率的定义是

$$\pi_{r,r}=\frac{\partial^2 E}{\partial \alpha_r^2}=\frac{\partial q_r}{\partial \alpha_r} \tag{12.11}$$

$\pi_{r,r}$ 是个负的数量，这表明当原子 r 的有效电负性增大($\delta\alpha_r<0$)时，q_r 增加。例如当芳烃中某个碳原子 r 受到亲电试剂 NO_2^+ 进攻时



$\longleftarrow NO_2^+(\delta\alpha_r<0,\delta q_r>0)$

负电荷要移向原子 r。相反，当 r 的有效电负性减少时 ($\delta\alpha_r>0$，例如当 r 受亲核试剂 NH_2^- 进攻时)，q_r 减少。$\pi_{r,r}$ 是原子 r 上的电荷密度对微扰敏感性(原子 r 上的电子极化率)的量度。$\pi_{r,r}$愈大，在某个试剂的微扰下，原子 r 上 π-电荷密度的变化(增大或减小)愈大。因此，对于 π-电荷作均匀分布的分子，无论在亲电或亲核反应中，总是 $\pi_{r,r}$ 最大的位置优先反应。例如萘的 α 和 β 位的原子自极化率分别是 0.443 和 0.405(以 $1/\beta$ 为单位)，这与萘不管是硝化或氨基化总是 α 位发生取代的事实相一致。

用 π-电荷密度或原子自极化率来预测反应性能，是假定试剂和受体间的静电力起主导作用，这当进攻试剂是带电荷的离子时是可行的。若试剂是电中性分子或自由基时，静电作用一般很小，这时候可用自由价 F_r 来预见反应活性。图 12-40 给出了芳烃的甲基取代速度和芳烃分子的最大自由价的关系(假定反应速度决定于分子的最大自由价)。图中纵坐标的 PRF 称为部分速度因子(partial rate factor)，它是指某位置上速度常数和作为标准的苯的每个位置上的速度常数之比。例如测得萘在 α 位的硝化速度常数为 K，苯的硝化速度为 K_0，萘有四个 α 位，苯有六个等同位置，故萘在 α 位的部分速度因子为

$$\mathrm{PRF}=\frac{K}{4}\Big/\frac{K_0}{6}$$

实验证明，用自由价来判断反应性能并不限于自由基反应。许多其它类型的反应，包括离子反应在内，往往也可用自由价来指示活性。

现在我们来看孤立分子近似中用电荷密度、极化率和自由价等来判断反应活性究竟有点什么理论根据。当试剂 R 向受体分子的原子 r 进攻时在试剂 R 的微扰之下，原子 r 附近的 π-电荷倾向于向 r 集中或离开 r (由 R 是亲电性或亲核性试剂而定)。微扰的结果是使 r 的库仑积分和键 r—s、r—t 的共振积分发生变化。如果把共轭体系的 π 电子能量看成是 α_r、β_{rs} 和 β_{rt} 的函数，按 Taylor 级数将此函数展开，得到微扰引起共轭体系的 π 电子能量增加为

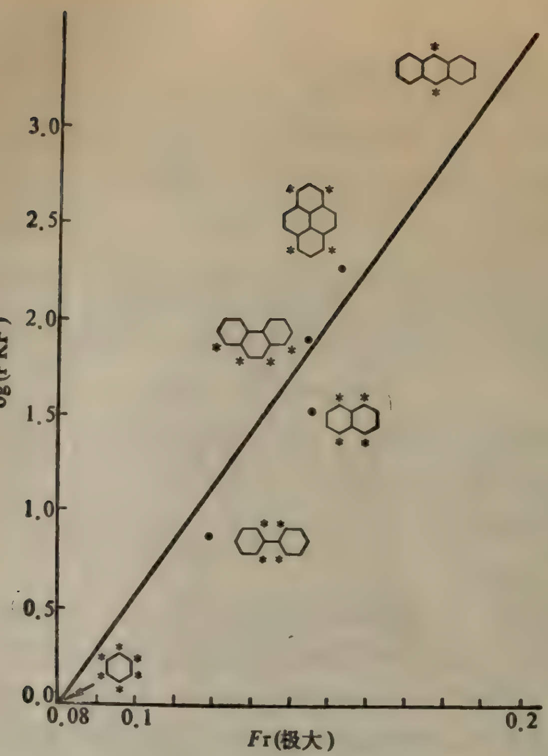


图 12-40　CH_3 取代速度和每个分子的最大自由价的关系图

(图中用星号标出最大自由价位置)PRF定义为$\frac{k}{m}\Big/\frac{k_0}{6}$，$k_0$ 是苯的速度常数，

6 是苯的 6 个等价位置，k 是具有 m 个最大自由价位置的速度常数。

$$\delta E_\pi=\frac{\partial E_\pi}{\partial\alpha_r}\delta\alpha_r+\frac{1}{2}\frac{\partial^2 E_\pi}{\partial\alpha_r^2}(\delta\alpha_r)^2+\frac{\partial E}{\partial\beta_{rs}}\delta\beta_{rs}+\frac{\partial E_\pi}{\partial\beta_{rt}}\delta\beta_{rt}+\cdots$$

(12.12)

$$=q_r\delta\alpha_r+\frac{1}{2}\pi_{r,r}(\delta\alpha_r)^2+2\,p_{rs}\delta\beta_{rs}+2\,p_{rt}\delta\beta_{rt}+\cdots$$

其中对α_r展开到二级微扰，并引用了 $\partial E_\pi/\partial\alpha_r=q_r$，$\partial^2E_\pi/\partial\alpha_r^2=\pi_{r,r}$，$\partial E_\pi/\partial\beta_{rs}=2\,p_{rs}$。设 $\delta\beta_{rs}=\delta\beta_{rt}=\delta\beta$，并注意 $p_{rs}+p_{rt}=\sum N_r=$常数$-F_r$，上式变为

$$\delta E_\pi=q_r\delta\alpha_r+\frac{1}{2}\pi_{r,r}(\delta\alpha_r)^2-2\,F_r\delta\beta+\text{常数}+\cdots \qquad (12.13)$$

当试剂R是离子时，长程的库仑力占优势，式(12.13)右边前两项起决定作用，其中第一项又比第二项重要。对电荷密度不均匀的分子，只要考虑第一项，即q_r决定其反应性能。在亲电反应中，$\delta\alpha_r<0$，q_r愈大，反应速度愈大。亲核反应中，情况正好与之相反。对电荷密度处相同的分子，反应活性的差别就决定于第二项，此时 δE_π 值只依赖于 $\pi_{r,r}$，而与 $\delta\alpha_r$ 的符号无关。无论反应是

亲电的还是亲核的，总是 $\pi_{r,r}$ 最大的位置反应速度最大。若试剂R是不带电荷的自由基，这时库仑力降居次要地位，决定 δE_{π} 的是(12.13)式右边第三项。随着自由基R·向受体分子接近，被攻击的原子 r 附近的 π 电子逐步和 R·键合，原子 r 有从共轭体系中分离出去的倾向，故r—s和r—t键的 π 键减弱，它们的键积分都从某个标准值 β(负值)向零方向变动，$\delta\beta>0$。因此，(12.13)式右边第三项是随 F_r 的增大而减低，即 F_r 愈大的位置活性愈大。

12.4.1.2 位能面[56]

量子理论在化学动力学和化学致癌作用研究中的最重要作用是计算一个反应 体 系 的 位能面。以下我们简略地介绍这种计算所依据的原理。 一个分子的总 Hamilton 是一些动能项和位能项之和，可以表示为

$$\hat{H}=\hat{T}_{n}+\hat{T}_{e}+\hat{V}_{ee}+\hat{V}_{en}+\hat{V}_{nn} \tag{12.14}$$

其中下标 e 和 n 分别指电子和核。

分子的总能量和其波函数由 Schrödinger 方程

$$\hat{H}\Psi=E\Psi \tag{12.15}$$

的解给出(这些已在第二章中较详尽地讨论过)。我们已经知道，式(12.15)对我们说来，是十分复杂的。不过，可通过引用Born-Oppenheimer近似予以简化。这样，便可以略去(12.14)式中的核动能项 $\hat{T}_{n}$，而得到电子 Hamilton $\hat{H}$。

$$\hat{H}_{e}=\hat{H}-\hat{T}_{n} \tag{12.16}$$

$\hat{H}_{e}$ 依赖于电子和核的坐标。例如 H_2 的 $\hat{H}_{e}$ 是

$$\hat{H}_{e}=-\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2)+\frac{e^2}{r_{12}}-\left(\frac{e^2}{r_{\alpha1}}+\frac{e^2}{r_{\beta1}}+\frac{e^2}{r_{\alpha2}}+\frac{e^2}{r_{\beta2}}\right)+\frac{e^2}{R}$$

当核构型被指定时，$\hat{H}_{e}$只是电子坐标的函数。对于每一个可能的核构型，解 Schrödinger 方程

$$\hat{H}_{e}\Psi_i^e=U_i\Psi_i^e \tag{12.17}$$

得到指定核构型下的电子波函数 Ψ_i^e(下标 i 标记电子态)和能量 U_i。U_i 包括电子能量和核间排斥能，所以称它为指定核构型时体系的位能。式(12.17)可用价键法或分子轨道法近似求解。当用分子轨道法求解时，可用半经验自洽场法或 ab initio 法解出指定核构型下的 Roothaan 方程，将得出的分子轨道组成反对称的 Slater 函数，就是近似的电子波函数 Ψ_i^e。由各分子轨道的轨道能、分子的电子组态以及核间排斥能 V_{nn} 可以求出 U_i 的近似值。对体系的每一个可能的核构型都作如上的计算，然后将 U_i 值对核构型作图。对 多 原 子分子可给出多维的位能面 (如图 12-41)。

图 12-41 是线型体系H_3的位能面。图中将位能函数$U(R_1, R_2)$的等值点连结成等高线以显示位能面。

位能面在化学致癌的量子生物学研究中获得了较广泛的应用。

12.4.1.3 分子静电势

在本书的第 2 章已经讨论过有关分子静电势的概念和计算原理。在这以后的一些章节中，我们具体介绍了分子静电势的应用实例。近几年来，分子静电势在化学致癌研究中应用日益增多，也取得了可喜的进展(参见 4.2.3 和 4.8)。

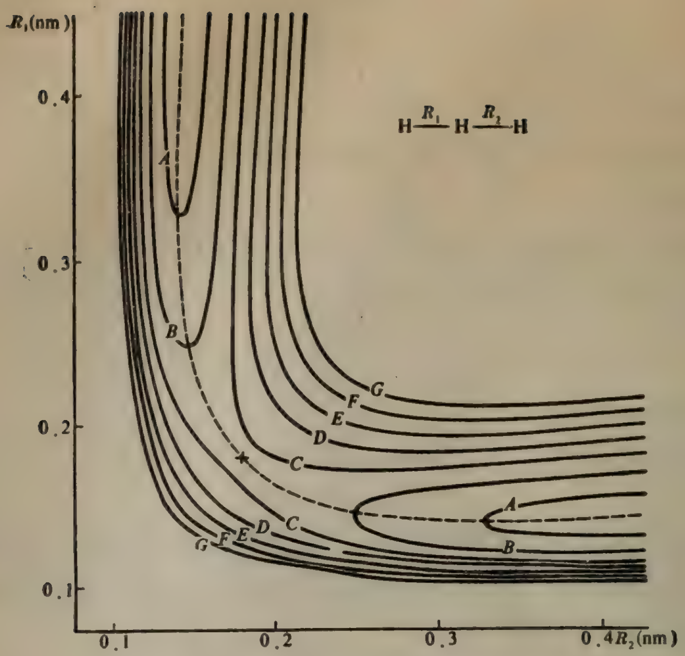


图 12-41 线型构型的 H_3 的位能等值线

等值线 A 的能值比 H+H_2(平衡核间距)的能量高 0.1 电子伏 (9.65 kJ/mol)。按 $A,B,C,D\cdots$ 顺序,每两条邻近的等值线能量相差 0.2 电子伏

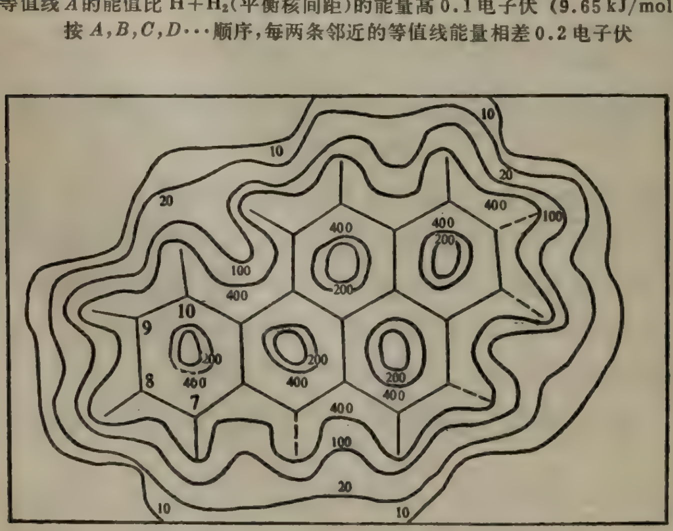


图 12-42 苯并[a]芘分子平面分子静电势等值线图

我们知道,分子静电势可以精确而直观地描绘出分子周围空间的静电作用性质,因而对于了解分子间的相互作用具有重要的作用。

Hariharan等[57]应用ab initio计算苯并[a]芘及其代谢物,3-甲基胆蒽及其代谢物的分子静电势,得到了分子静电势等值线图(图 12-42 至 12-44)。

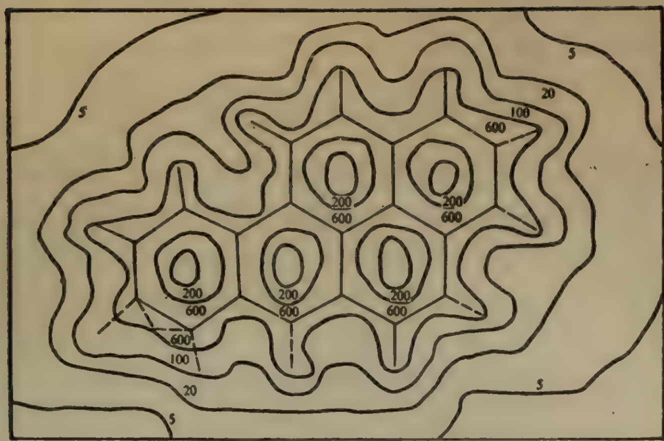


图 12-43　苯并[a]芘 7,8-氧化物分子平面分子静电势等值线图

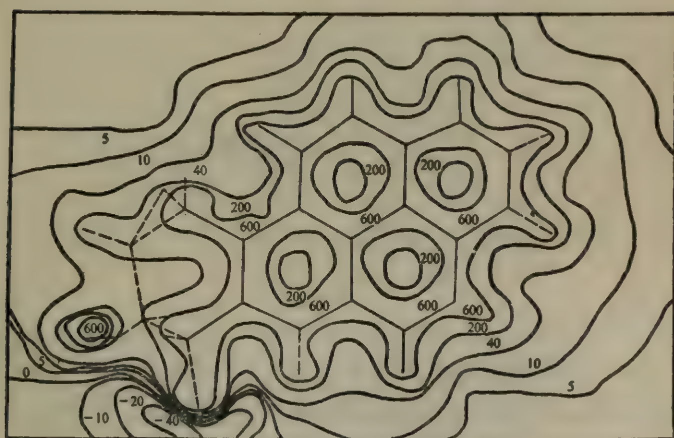


图 12-44　苯并[a]芘 7,8-二氢二醇-9,10-环氧化物 $A\alpha,\beta,\alpha$ 分子平面分子静电势等值线图

从这些分子静电势等值线图可以看到，它们与实验结果是一致的。这些图形象地向我们指示出C 10 是亲电性攻击的主要位置。苯并[a]芘经代谢生成 7,8-二氢二醇-9,10-环氧化物后，分子体系被更大的正电势的套所包围，就整体而言，它是一个带正电的实体，由此可以推测其靶分子应带一定的负电荷。当它们由于静电吸引而接近时，苯并[a]芘 7,8-二氢二醇-9,10-环氧化物的分子静电势特征的细节，就将被靶分子进一步感知和识别。

从图 12-42 至 12-44 的分析还可以看出，这些图还可以用作“亲电性”的指标。从图 12-44 可知，苯并[a]芘 7,8-二氢二醇-9,10-环氧化物进一步将导致环氧桥在 C 10 处断开，氧仍与 C 9 相连结，如 $-\mathrm{O}-\mathrm{C}\,9-{}^{+}\mathrm{C}\,10$。

从 4.2.3 和 4.8 已经知道，碱基、核苷、核苷酸、单螺旋、双螺旋或变性 DNA 的环氮 N 和羰

基氧原子的分子静电势深度逐步增加。实验表明，跟这种变化相一致的是它们与亲电性致癌物质的反应性也随静电势深度增大而增加。

12.4.2 多元回归分析

寻找一种事物同其它事物的相关关系的方法叫多元回归分析。几乎在环境化学的每个领域(包括化学致癌的研究)都能找到应用实例，但预报仅只是大致的估计值。线性方程系数正负表示正、负相关；系数权重比较表示这些变量在相关关系中贡献的大小。

算法为：若因变量为Y_n，自变量为$X_{1n}, X_{2n}, \cdots, X_{mn}$，则令

$$L_{ij}=\sum_{k=1}^{n}(X_{ik}-\overline{X}_i)(X_{jk}-\overline{X}_j) \quad i,j=1,2,\cdots,m \tag{12.18}$$

$$L_{io}=\sum_{k=1}^{n}(X_{ik}-\overline{X}_i)(Y_k-\overline{Y}) \quad i=1,2,\cdots,m \tag{12.19}$$

$$L_{oo}=\sum_{k=1}^{n}(Y_k-\overline{Y})^2 \tag{12.20}$$

其中

$$\overline{X}_i=\frac{1}{n}\sum_{k=1}^{n}X_{ik} \quad i=1,2,\cdots,m$$

$$\overline{Y}=\frac{1}{n}\sum_{k=1}^{n}Y_k$$

此时，回归方程为：

$$Y=a+b_1X_1+b_2X_2+\cdots+b_mX_m \tag{12.21}$$

根据最小二乘法原理，系数 $b_1, b_2, \cdots, b_m$ 必需满足如下多元线性方程组：

$$\begin{aligned} l_{11}b_1+l_{12}b_2+\cdots+l_{1m}b_m&=l_{10}\\ l_{21}b_1+l_{22}b_2+\cdots+l_{2m}b_m&=l_{20}\\ \cdots&\\ l_{m1}b_1+l_{m2}b_2+\cdots+l_{mm}b_m&=l_{mo} \end{aligned} \tag{12.22}$$

常数项为：

$$a=\overline{Y}-\sum_{i=1}^{m}b_iX_i$$

先输入样本数N，自变量数M及自变量、因变量X_{ik}和Y_k。造线性方程组。用 Gauss 全主元消去法求解。算出系数 a 和 b_i，线性方程相关系数 R 和剩余标准离差 S。R用于衡量回归效果；S 表达回归方程精度。

作为统计规律，样本数愈多效果愈好，一般N比M应大于 5。再者，自变量不应有较大测量误差。但实际上化学致癌研究所碰到的量本身误差就较大，难以用上述方法作多元回归处理。

有人建议用几何平均函数回归法，可以估计到自变量测量误差较大的情况，使预报更为准确。再就是，相关性并不等于必然性。不要看见甲与乙相关性较好，就肯定乙一定决定甲。可能还有丙通过乙决定甲。总之，多元回归分析简单、直观、常用，但诸变量关系往往不是简单的线性方程所能描绘，故而较为粗糙，防止滥用。

化学致癌过程是一个非常复杂的过程，为了保证结果的可靠性，一个有益的办法是时刻注意统计观点。因此，对每一个系列，我们总是力图取 50 个以上或更多的化学样本以开始我们的结构致癌活性关系处理，从统计观点讲，这是一个可靠的数目[47]。

甚至最简单的最小二乘拟合往往也会得到若干有启发性的结果。图 12-45 是 36 个有实验数据的具湾区或狭湾区的PAH，其湾区ΔE与其致癌活性（五级相应取作 1,2,3,4 和 5）的最小二乘结果（表 12-21 是图 12-45 的数据表，化合物名称皆用缩写）。

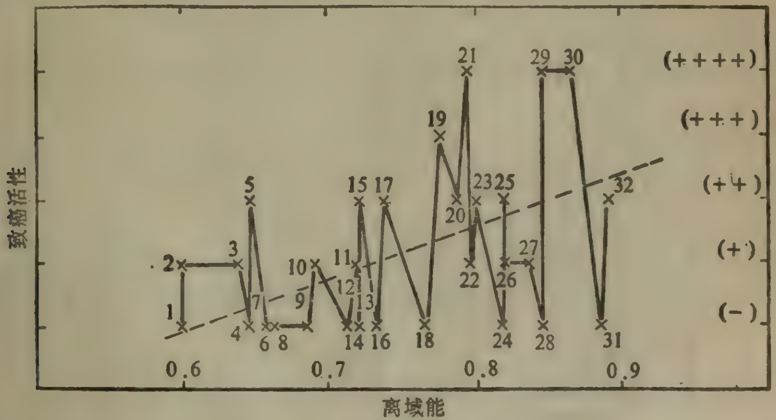


图 12-45

从图 12-45 可见，其相关性极差，相关系数只有 0.4092，致癌活性沿回归线两侧剧烈地波动。具“异常”高的致癌活性的 PAH 30,29,21，两活性中心的距离为适当的 0.278 nm，而具“异常”低活性的PAH 31,28,24 则有几个脱毒的L区，且两活性中心的距离很大[58]（参见 12.2.1.4）。

12.4.3 模式识别

12.4.3.1 模式识别法简介

模式识别是人工智能的一个分支，目的是弄清判别过程的机制，并用计算机使之自动进行。模式识别基本原理是“物以类聚”。

模式识别分为五步：

1．收集原始数据

要尽可能收集那些对化学致癌物和致癌过程有潜在性影响的指标，过多又会增加无益的计算量。样本数应大于维数五倍，分类预报效果方好。

2．数据表达和整理

某些指标不是数，是概念，如午、夜、黄昏、阴、晴、多云、清、浊、混、强致癌、致癌、微弱致癌等等。有的甚至是模糊概念，则应建立标度，转化为计算机能懂的数码。样本矩阵要归一化，以克

表 12-21

NO	化合物	能量	指数
1	N12aA	0.6	1
2	BoPH	0.6	2
3	C	0.639	2
4	BbC	0.647	1
5	BcC	0.648	3
6	PH	0.658	1
7	PI	0.662	1
8	TR	0.664	1
9	PH_9 10bTR	0.687	1
10	DBfgopNC	0.692	2
11	BeP	0.714	1
12	BgC	0.719	2
13	N12bTR	0.721	1
14	BbTR	0.722	1
15	DBajA	0.722	3
16	BaPP	0.733	1
17	DBahA	0.738	3
18	BaA	0.766	1
19	N1234defC	0.775	4
20	DBacNC	0.787	3
21	BaP	0.794	5
22	TBacjC	0.795	2
23	DBdefpC	0.8	3
24	DBajNC	0.818	1
25	DBhrstpp	0.819	3
26	BaN218hijNC	0.82	2
27	BaN812cdeNC	0.837	2
28	BaNC	0.845	1
29	DBbdefC	0.845	5
30	BrstPP	0.866	5
31	DBalPC	0.885	1
32	TBbdefpC	0.89	3

$R=4.092$,　A=6.348,　B=2.617,　$S=1.1575$

服量纲影响；加权以增加类间分离的可分离性。

3．特征选择

从样本的诸特征中选择对问题有重要影响的特征的方法叫特征选择。原始N维空间降为维数少的M维空间，还应保证原点群空间分布在降维时其类别可分离性不应有较大影响，故又叫映照。降维使问题约化，计算量减少；删除了次要或无关因素，使问题清晰化。

特征选择可用主成份分析法进行。取少数特征值较大的特征函数建立新的维数空间。映照仅限于正交归一线性变换，原分布不会走样。

特征选择也可用分离指数法进行。下式中分离指数 P 为类别可分离性之量度。可将诸特

征依次代入，筛选变量，降低维数，取分离指数降低为最小者，即为最佳线性映照。

$$P=\frac{\sum_{k=1}^{m}(\bar{u}_k-\bar{v}_k)^2}{\left(\frac{1}{L_1}\sum_{i=1}^{L_1}\sqrt{\sum_{k=1}^{m}(u_{ik}-\bar{u}_k)^2}\right)\left(\frac{1}{L_2}\sum_{i=1}^{L_2}\sqrt{\sum_{k=1}^{m}(v_{ik}-\bar{v}_k)^2}\right)} \tag{12.23}$$

其中，m为维数，L_1、L_2为两类点数；u_{ik}、v_{ik}为两类点中 i 点坐标；$\bar{u}_k$、$\bar{v}_k$为两类重心。

4．分类判别

有最短距离法，设全部训练点集分属 G 类，各类样本点聚于重心周围，则任意点 i 与各重心距离最短者的类就是 i 点应归属的类。首先求出 G 类训练点重心$P_{g1},P_{g2},\cdots,P_{gk}$。

$$P_{gk}=\frac{1}{n_g}\sum_{i=1}^{n_g}X_{gki}(g=1,2,\cdots,G;k=1,2,\cdots,m) \tag{12.24}$$

式中 m 是维数，n_g 为第 g 类训练点数，X_{gki}与第 g 类训练点坐标。点 i 与 G 个重心之欧氏距离为

$$d_i,P_g=\sqrt{\sum_{k=1}^{m}(X_{ik}-P_{gk})^2}\quad(g=1,2,\cdots,G)$$

若满足下式则 i 点属于 g^*类。

还有线性函数判别法。若有线性函数存在于 R 类之间，

$$g_i(x)=W_{i0}+\sum_{k=1}^{m}W_{ik}X_k \tag{12.25}$$

当点 x 属于 i 类时，则有

$$g_i(x)>g_j(x)\quad(i,j=1,2,\cdots,R)$$

则称 $g_i(x)$为线性判别函数。为方便计，令 $W_i(W_{i0},W_{i1},W_{i2},\cdots,W_{im})$ 为第 i 类权重向量；$Y(1,X_1,X_2,\cdots,X_m)$为增广模式向量，则线性判别式减化为

$$g_i(Y)=W_i\cdot Y\quad(i=1,2,\cdots,R) \tag{12.26}$$

在两类分类问题中，$R=2$，则

$$g_1(Y)-g_2(Y)=(W_1-W_2)\cdot Y=W\cdot Y$$

若 $W\cdot Y>0$，则 Y 属于 W_1 类；$W\cdot Y<0$，则 Y 属于 W_2 类；$W\cdot Y=0$，是类间界面方程。可先求出二组训练点重心，以二重心中垂面之法矢量为 W 之初始值，采用错分点反馈修正法逐步修正，直至权重向量W满足最大限度分类为止。

最常用的是非线性映照法。在尽可能保持样本点间距离不变条件下，把M维空间映照到二维空间，在可显示的二维空间内据其分布实现样本分类和识别。对于样本矩阵$X_{ik}(i=1,2,\cdots,n;k=1,2,\cdots,m)$，定义降维产生误差的误差函数为

$$E=\frac{1}{\sum_{i<j}^{n}d_{ij}^*}\sum_{i<j}^{n}\frac{(d_{ij}^*-d_{ij})^2}{d_{ij}^*}\quad(i,j=1,2,\cdots,n) \tag{12.27}$$

d_{ij}^*为 M 维空间内 i,j 点距离；d_{ij} 为映照后二维空间内 i,j点距离，可依下式计算

$$d_{ij}^* = \sqrt{\sum_{k=1}^{m}(X_{ik}-X_{jk})^2} \quad (i,j=1,2,\cdots,n) \tag{12.28}$$

$$d_{ij} = \sqrt{(Y_{i1}-Y_{j1})^2+(Y_{i2}-Y_{j2})^2} \quad (i,j=1,2,\cdots,n) \tag{12.29}$$

(12.28) 式中，d_{ij}^*为常数，但 d_{ij}中包含 2 N个未知数$Y_{i1},Y_{i2}(i=1,2,\cdots,n)$。初值选用主成分分析的前两个主成分组成初始二维空间，然后按最优化方法使误差函数为极小，用 Newton 迭代法求解。

5. 显示与预报

在 XY 二维绘图仪上给出样本点分布和类间分界线。标出未知点在图上的位置，点在哪个类里就归属哪类，从而达到预报目的。

12.4.3.2 模式识别在化学致癌作用研究中的应用实例

我们知道，N-亚硝基化合物由于其在预防环境癌和职业癌方面的重要意义，已积累了近200个化合物的结构致癌活性关系的实验资料，其中数据比较可靠并能据以进行按五级致癌标度进行分级的达 153 个。但是，其结构致癌性关系，尤其是定量结构致癌性关系，仍然是一个没有解决的问题；对其致癌机理的研究也还存在着不同的观点。[59,60]

我国科学工作者戴乾圜等[61]应用多元回归的定量模式识别方法，在计算机上对上述 153 个N-亚硝基化合物的结构致癌活性关系的资料进行了处理，从大量可能有意义的描述项中筛选出了10个描述项(表12-22)。这10个描述项清楚地表明，N-亚硝基化合物致癌的必要条件同样要在代谢过程中产生两个活泼的亲电中心，其中一个反应中心是传统观点的经酶催化或非酶作用产生重氮阳离子的 α-位；另一个反应中心则是在脂肪链上羟化后，通过硫基转移酶、磷基转移酶或乙酰基转移酶等所产生的活性酯。后一反应中心的最有利位置是 β-位或者是正好与ω-位相当的 γ 位。α-位与β-位或 γ 位之间发生双官能烷化的有效距离恰好接近0.280—0.300nm。

化学致癌作用的量子生物学研究提示我们，恰当地引入模式识别的优点，就是可以客观地从结构致癌活性关系的实验事实，鉴别若干起决定作用的因素。当然，模式识别法在化学致癌的其它研究领域也具有这样的优点。

从表12-22可以看出，N-亚硝基化合物的α-位以及 β-位间的双官能烷化作用，从两者较大的正相关系数和方差比可以判明，对N-亚硝基化合物的致癌作用起着重要的贡献。在环上作为ω-位出现的γ-位，当其有利于亲电反应的邻近基团协助可能时，表现出高度的正相关。戴乾圜等认为，这显然与双区理论一致，α-碳链与β-碳末端的距离，以及以锯齿形存在的 α-与 γ 碳链末端的距离，皆趋近于 DNA 互补碱基对两负性中心的距离(即0.280—0.300nm)。这表明 N-亚硝基化合物诱发致癌作用的关键步骤，也是 DNA 互补碱基对的横向交联[6,41,58]。所以，通过N-亚硝基化合物的模式识别处理，得到的 N-亚硝基化合物致癌机理与双区理论是完全一致的。上述特定的双官能烷化概念，使得 N-亚硝基化合物的结构致癌活性关系皆可以相应的化学性质或物理性质来加以解释。根据上述10个描述项的模式识别，计算与实验的符合率达 97.3%[61]。

必须指出，戴乾圜等是在通过对苯系母体PAH、取代的PAH、非交变PAH等大多数化学致癌

表 12-22 N-亚硝基化合物SCAR模式辨认的描述项

序号	名称	选取方法	涵义	回归系数	方差比
1	α-H积	两个α-碳上氢原子数之积	体现重氮盐产生的速率和位阻	0.548	1.139
2	β-H积	氢原子数最多的两个β-碳上氢原子数之和	体现β-位活性酯产生的几率及位阻	0.436	0.452
3	β-位促活	β-位有—Cl, —Br, —OSO_3H等=3 β-位有—OH,—C═C等=1；否则为0	β-位这类基团的存在将有利于缩短代谢活化过程，有利于横向交联	0.610	0.178
4	γ-位邻助	O、N、S等通过三环成五环过渡态对γ-位的邻助效应，有=1；无=0	γ-位通过这类作用将成为有效的第二活性区	1.340	0.158
5	α,β-位取代基校正	一侧α-碳有O或N相连时=另一侧α, β-H之和，两侧α-位皆有O或N时=−4 β-H皆被F、N等原子取代=−4	可能经无酶作用而裂解； 产生的重氮盐不稳定将失活； β-位被阻断将失活	0.569	0.921
6	修正碳原子数	取值=2+|NC−4| NC为碳原子总数	模拟酶活性中心空隙而出现的影响	−0.119	0.082
7	脱毒基团	对—CH_2OH,—C═CH_2,—$ArCH_3$取1；对—CHO,—COOH取3	表达其具有成羧酸脱毒及活化的双重性，主导趋势为脱毒	−0.458	0.093
8	环系效应	六员环以下取−1，六员环取0，六员环以上取1，环上每一取代甲基增加−0.1	模拟环因Baeyer张力和Pitzer张力及因环大小不同对β-碳正离子的形成和稳定性的影响	0.705	0.056
9	修正N数	分子中不与亚硝基相连的氮原子数	有利于亲特定器官的性质而增强活性	0.820	0.119
10	类芳环取代效应	两侧均无类芳基体积的基团取0，一侧有此类基团取另一侧α-H数	模拟酶与底物作用的锁钥效应	0.445	0.102

剂系列进行系统的量子化学计算的基础上，提出致癌作用的双区理论，从而建立相应的定量关系式，并加以推广的。他们在这些工作的基础上，借模式识别，建立了N-亚硝基化合物的结构致癌活性关系的双区理论模型。其中，有关模式识别处理的数学方法，请参见文献[61]。

综上所述，可以认为，量子生物学在探讨化学致癌物分子的电子结构与致癌作用的关系以及致癌机理等方面是卓有成效的，但毕竟还只是初步的，不完整的。当然，正如Scribner所认为的那样，分子轨道法在化学致癌的研究中被一些人所滥用的情况也是应当引起注意的[17]。总之，迄今为止，有关化学致癌作用的机制仍未从根本上完全阐明，许多研究工作还仅仅是开始。但是，随着量子生物学这门新兴的边缘学科的发展，随着其它关系到化学致癌作用研究的学科的发展，化学致癌作用机制的揭示和癌症的防治必将会以更大的速度取得新的进展。

参考文献

[1] Schmidt, O.: Z. Physik. Chem., B 42:83 (1939).
[2] Pullman, A. et al.: Rev. Sci., 3:117 (1946). Adv. Cancer Res., 3:117 (1955).
[3] Jerina, D. M. et al.: "Drug Metabolism from Microbe to Man", Parke, D. V. et al. (Eds.), (Taylor and Francis, London), (1976).

[4] Jerina, D. M. et al.: "In Vitro Metabolic Activation in Mutagenesis Testing", de Serres, F. J. et al. (Eds.), (Elsevier, Amsterdam), (1976).
[5] Jerina, D. M. et al.: "Microsome and Drug Oxidations", Ullrich, V. et al. (Eds.), (Pergamon Press, Oxford), (1977).
[6] 戴乾圜：中国科学，(10):964 (1979).
[7] 引自 Shimkin, M. B. et al.: Prog. Exp. Tumor Res., 11:1 (1969).
[8] Yamagiwa, K. et al.: J. Cancer Res., 3:1 (1918).
[9] Tsutsui, H.: Cann, 12:17 (1918).
[10] Passey, R. D.: Brit. Med. J., 11:1112 (1922).
[11] 程元凯编：«致癌性多环芳烃»，人民卫生出版社，(1980).
[12] 戴乾圜：环境化学，3:1 (1984).
[13] Kootstra, A.: Nucleic Acids Res., 10:2775 (1982).
[14] Sanders, C. A.: Biochem. Biophys. Res. Commun., 78:1034 (1977).
[15] Sculley, T. B. et al.: Cacer Res., 43:1688 (1983).
[16] Macleod, M. C. et al.: Proc. Natl. Acad. Sci. U. S. A., 77:6396 (1980).
[17] Scribner, J. D.: J. Natl. Cancer Inst., 55:1035 (1975).
[18] Knop, J. V. et al.: Int. J. Quantum Chem., 23:713 (1983).
[19] 戴乾圜：中国科学(B辑)，(8):716 (1983).
[20] Jerina, D. M. et al.: Adv. Pharm. Toxi., 9:53 (1979).
[21] Kapitulnik, J. et al.: Cancer Res., 38:354 (1978).
[22] Levin, W. et al.: Cancer Res,. 38:1705 (1978).
[23] Whalen, D. L. et al.: J. Am. Chem. Soc., 100:5218 (1978).
[24] Buening, M. K. et al.: Cancer Res., 39:5063 (1979).
[25] Wislocki, P. G. et al.: J. Natl. Cancer Inst., 63:201 (1979).
[26] Wood, A. W. et al.: Cancer Res., 40:2876 (1980).
[27] Levin, W. et al.: Cancer Res., 40:3910 (1980).
[28] Wood, A. W. et al.: Cancer Res., 40:1985 (1980).
[29] Slaga, T. J. et al.: Cancer Res., 37:4130 (1977).
[30] Arcos, J. C. et al.: Adv. Cancer Res., 11:305 (1968).
[31] Jerina, D. M. et al.: Pure Appl. Chem., 30:1033 (1978).
[32] 白迺彬等：分子科学学报，(1):97 (1982).
[33] 陈凯先、吴吉安、嵇汝运：分子科学学报，(3):1 (1983).
[34] 白迺彬：生物化学与生物物理进展，(5):1 (1980).
[35] Heerzberg, G.: Proc. Roy. Soc., A 262:291 (1961).
[36] Lathan, W. A. et al.: J. Am. Chem. Soc., 93:6377 (1971).
[37] 袁云程：化学通报，(3):267 (1979).
[38] 曾广植，化学通报，(7):404 (1980).
[39] 曾昭氘等：化学通报，(4):62 (1984).
[40] 陈念陔等：分子科学学报，1:123 (1981).
[41] 戴乾圜：多环芳烃致癌性能的定量分子轨道模型——双区理论的进展和应用，«环境化学论文集»，科学出版社，(1982).
[42] Memory, J. D.: Inter. J. Quantum Chem., 15:363 (1979).
[43] Weinstein, I. B. et al.: Science, 193:592 (1976).
[44] 陈洪等：北京轻工业学院学报，2:1 (1984).
[45] Swenson, D. H. et al.: Biochem. Biophys. Res. Commun., 60:1036 (1974).
[46] Ln, J. K. et al.: Cancer Res., 37:4430 (1977).
[47] 陈洪等：环境化学，1:253 (1982).
[48] 戴乾圜等：科学通报，27:696 (1982).
[49] 王连生等：环境科学学报，7:240 (1987).
[50] Lavery, R. et al.: Inter. J. Quantum Chem., QBS., 5:21 (1978).

[51] Pullman, B. et al.: Inter. J. Quantum Chem., QBS., 16:175 (1979).
[52] Sarma, D. S. et al.: "Cancer, a Comprehensive Treatise", Becker, F. F. (Ed.), (Plenum Press), (1975).
[53] Fredrickson, D. S. 等著，陈来成译：«癌的病因与防治»，科学普及出版社，(1981).
[54] Stezowski, J. J. et al.: Cancer Res., 44:5555 (1984).
[55] 温元凯、邵俊：大自然探索，(1):85 (1984).
[56] 朱永等编著：«量子有机化学»(下册)，上海科学技术出版社，(1986).
[57] Hariharan, P. C. et al.: Inter. J. Quantum Chem., QBS., (6):223 (1979).
[58] 戴乾圜：环境化学，3:1 (1984).
[59] Magee, P. et al.: Biochem. J., 83:114 (1962).
[60] Lijinsky, W.: "Genotoxicology of N-Nitroso Compounds", Rao, T. K. et al. (Eds.), (Plenum Press, New York), (1984).
[61] 戴乾圜：环境化学，6:1 (1987).

刘次全
白适彬
陈　洪

到期 91.4.3.
来源 王府井新
书价 7.20
单据号 600844
91.4.2.

中科院植物所图书馆
S0014731

58.1713
191

注　　意

1 借书到期请即送还。
2 请勿在书上批改圈点，折角。
3 借去图书如有污损遗失等情形须照章赔偿。

25650

京卡0761

ISBN 7-04-003216-3/Q·172　　定价7.20元